Journal of
Clinical Psychology

JULY 1951
Volume VII Number 3

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Medical College, University of Vermont

JERRY W. CARTER, JR.
U. S. Public Health Service

HOWARD F. HUNT
University of Chicago

ELAINE F. KINDER
Rockland (N.Y.) State Hospital

SAMUEL W. HAMILTON

WILLIAM A. HUNT
Northwestern University

C. M. LOUTTIT
University of Illinois

ROBERT I. WATSON
Washington University

ELEANOR C. THORNE, *Business Manager*

Journal of Clinical Psychology
5 Pearl Street - - Brandon, Vermont

The Brown Schools

For Exceptional Children
RESIDENT TREATMENT CENTER

Six Resident Centers in and near Austin, Texas
Administrative Offices 1110 East 32nd Street, Austin

Bert P. Brown, F.A.A.M.D., *President*

Paul L. White, M.D., F.A.C.P.,
Certified by American Board of Psychiatry

Ernestine B. Blackwell, Ed.D.
Psychological and Educational Director

Jesse J. Villarreal, Ph.D.
Speech Pathologist

O. Hoyt Williams, M.A.
Director of Vocational Education

Summer Camp: June - August

Resident Educational Program:
September - May

Lodges on Lake Travis

Private Spring-fed Pool

Physicians
Psychologists
Parent Consultants
Nurses
Counselors
Teachers in Special Education
Occupational Therapists
Recreational Therapists
Teachers in Vocational Education

- **Exceptional Children Ages 4 to adulthood**
- **Comprehensive medical, psychiatric, psychologic and educational evaluation**
- **Homogeneous grouping**
- **Controlled environment**
- **Milieu approach to development of the whole child**
- **Specialized therapies**
- **Vocational training**

All the disciplines of physical and mental health in a coordinated team approach work together for the continued growth of the exceptional child

---

Viewbook on request

Interviews arranged with the Consulting Staff at the Schools or in your home

● An exclusive school for children with emotional or educational difficulties

JOURNAL OF CLINICAL PSYCHOLOGY

VOL. VII JULY 1951 No. 3

CONTENTS

On the Disuse and Misuse of P, Q, Qs and O Techniques in Clinical Psychology. R. B. Cattell, University of Illinois 203

Research in Clinical Psychology: 1950. William Schofield, University of Minnesota 215

The Prognostic Value of the Metenym Test in a Follow-up Study of Psychosurgery Patients and their Controls. Joseph Zubin and Charles Windle, Columbia University 221

A Study of the Relationship between the Verb-Adjective Quotient and the Rorschach Experience Balance. William Hays, Saul Gellerman and William Sloan, Lincoln State School and Colony 224

The Figure-Background Relationship in Children with Cerebral Palsy. Jane E. Dolphin, Florida State University, and William M. Cruickshank, Syracuse University 228

A Factor Analysis of Some Szondi Pictures. Wilson H. Guertin, Ionia State Hospital 232

Sex Differences in the Rosenzweig P-F Study, Children's Form. George Spache, University of Florida 235

Symbolic Meanings in the Rorschach Cards: A Statistical Study. Ephraim Rosen, University of Minnesota 239

The Relationship between Certain Measurable Functions of Autonomic Nervous System Activity and Color Responses on the Rorschach Test. Halcyon Hughes, Leon J. Epstein, and Hudson Jost, University of Tennessee 244

A Behavior Rating Scale Suitable for Use in Mental Hospitals. Rubel J. Lucero and Bill T. Meyer, Fergus Falls State Hospital 250

An Objective Method of Evaluating Mental Status. John T. Rowell, Florida State University 255

An Analysis of the Elgin Prognostic Scale. Maurice Lorr, Phyllis Wittman and W. Schanberger, Elgin State Hospital 260

The Use of Tolserol (Myanesin) in Psychological Testing. Margaret Mercer and Arthur O. Hecker, Veterans Administration Hospital, Coatesville, Penna 263

A Comparison of the Performance of a Randomly Selected College Population on the MMPI and the P-S Experience Blank. Paul W. Cauffiel and William U. Snyder, Pennsylvania State College 267

The Clinician and his Predictions. Charles A. Dailey, St. Louis, Mo. 270

Directions for Administration of the Achromatic-Chromatic H-T-P. John N. Buck, Lynchburg State Colony 274

(Over)

CONTENTS—Continued

A Tabulation Method for Analysing Combinations of Rorschach Scores. Kenneth R. Hammond, University of Colorado 276

Use of the Blacky Pictures with a Child Whose Oedipal Desires are Close to Consciousness. H. Michal-Smith, Emanuel Hammer, and Herman Spitz, New Jersey State Diagnostic Center 280

The Application of Hypnosis to Nondirective Psychotherapy. Milton V. Kline, Westchester County Department of Health 283

A Note on Extended Findings with the MMPI in Predicting Response to Electro-Convulsive Therapy. John S. Pearson and Wendell M. Swenson, Minnesota State Division of Public Institutions 288

Norms for Scatter Analysis on the Wechsler Intelligence Scale. Benjamin Alimena, Fordham University and Catholic Charities Guidance Clinic 289

Editorial Opinions 291

Letter to the Editor 293

Book Reviews 294

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Special sections in each issue will be devoted to news, comments, and letters to the editors on matters pertinent to the profession. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

ORIGINAL ARTICLES are published only with the understanding that they are contributed exclusively to the *Journal of Clinical Psychology*. The publishers are not responsible for statements made or opinions expressed by contributors in articles printed in these columns. Copyrights cover publication in the *Journal of Clinical Psychology* and articles may not be reproduced without permission of the publishers.

MANUSCRIPTS must be typewritten, double-spaced with good margins, on one side of the paper only. The publishers will not be responsible for manuscripts lost in the mails or through unavoidable accident. Retain a carbon copy after sending the original to the editor. It is requested that all articles be as brief as consistent with clarity. All manuscripts are subject to editorial modification.

THE STYLE of all manuscripts including bibliographies must be in accordance with the instructions given in: Anderson, J. E., and Valentine, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.*, 1944, 41, 345-376. References should be arranged alphabetically according to the directions given above. Full address of the author should appear somewhere in the article, preferably at the end.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Rutland, Vermont, by Frederick C. Thorne, M.D. Subscription price $6.50; Canadian subscriptions $7.00; foreign subscriptions $7.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Reentered at the post office at Brandon, Vermont. Additional entry at the post office at Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1951, Frederick C. Thorne, M. D.

ON THE DISUSE AND MISUSE OF P, Q, Qs AND O TECHNIQUES IN CLINICAL PSYCHOLOGY

R. B. CATTELL

University of Illinois

Factor Analysis

Methods of scientific investigation are broadly of two kinds: experimental and statistical. Clinical psychology can make little use of the former, because there is so little of importance in the life of a human being that can be arbitrarily manipulated, and it must consequently fall back upon statistical analysis of life events *in situ*. This has been well brought out in Zubin's recent symposium [28]. The so-called "clinical method" of investigation is strictly a branch of statistical method, correlating life events and observations in the consulting room in a single whole.

Unfortunately, the clinician has been extraordinarily slow to recognize this generalization. Had he been quicker he could by now have availed himself of many powerful clinico-statistical tools. Throughout the Nineteenth Century physicians of unusual powers of observation and memory recognized a number of new disease syndromes by the feat of performing cluster analysis of symptoms without benefit of calculating machines. They also established causal, functional connections without using a T test to show that certain sequences of events occurred with a frequency beyond chance expectation. But in principle they were using, as all clinical method does, essentially the statistical methods we now explicitly recognize as such. The available methods have recently been very comprehensively summarized by Mensh [19], but without personal evaluation of the relative effectiveness of the methods or of the mutual criticisms by the enthusiasts for particular methods. The present article maintains a definite position: that the most fruitful and economical approach for the clinician at present is through factor analysis and especially through the intra-individual analysis methods which have been called O– and P–techinques [4].

Apart from specifically clinical problems, and in relation to general problems in social and biological sciences it can at least be said that the most powerful statistical method yet developed for recognizing underlying law and order in a wide array of variables observed *in situ* is factor analysis. Among clinical observations there is also no question that it offers possibilities of proving the number and nature of functional unities in an unstructured array of dynamic manifestations which far exceeds anything possible to unaided intuition. But the number of psychiatrists and clinical psychologists who have recognized this fact, to the extent of checking their theorizing and intuitive speculations by reference to factorial analysis of observations is probably less than one percent. Thousands glibly use such concepts as "ego strength"; but if the ego exists as a dynamic unity, varying in strength from person to person, then it should appear as a single factor when we correlate, by R-technique, measures of the variables or symptoms which allegedly manifest ego strength or weakness. If it does not so appear there is no scientific justification for continuing this clinical hunch. Actually, the C factor found in behavior ratings, subjective questionnaire responses and objective tests [4, 22] gives the ego concept the scientific support and clarification which it needs. This is but one illustration of the general theoretical position that if the fine intuitions of the clinician are to become established as part of a behavioral science, permitting calculation and testable prediction, they must prove themselves in the realm of experiment or of statistical inference. Psychoanalysis will become a science when it becomes a branch of factor analysis.

Description of P–, Q–, Qs–, and O–

Of the various factor analytic techniques, R–technique is perhaps least adapted to the clinician concerned with the individual case. It reveals and measures *common* traits, such as ego strength and the various functionally unitary drives which all

people possess in recognizably the same pattern but in different strengths. The clinician, however, wishes to go beyond this to the peculiar fixations of the individual case, and here the method of R–technique offers general theory but P–technique guides the diagnosis and therapy of the individual case. It contributes to general theory too in so far as the peculiarities of the individual case add a new page to theory.

In P–technique a set of measures is made upon the patient every day, on the same variables, before diagnosis and during therapy, extending perhaps over a hundred days or more. Correlatable series are thus created and every possible correlation among the variables can be worked out to provide a matrix for a factor analysis. The variables will consist largely of dynamic measures, including symptom strengths, interests in particular activities, dream and free association category frequencies, and attachments to particular individuals. But it is desirable in a complete design to include among them also measures of (a) the primary personality factors, notably the factors which measure degree of general neuroticism [22, 27], (b) the strength of various environmental presses important in the patient's real life adjustment and (c) some physiological variables. The importance of bridging across to R–technique, as prescribed in the first methodological recommendation, will come out in later discussion.

The correlation coefficient can show far more sensitively than the clinician's intuition alone which of these variables are responsively varying together or are showing that invariable sequence which we call causal connection. Thus P–technique is capable of linking particular symptoms with various personality source traits and ergs (drives), as has been shown in two demonstration cases [5, 9]. Indeed, the ergic structure behind particular dynamic manifestations is revealed with a certainty of qualitative (drive quality) analysis at least equal to that obtainable by free association, "projective" techniques or similar primitive clinical methods, and it does so while providing a quantitative basis for prediction which far exceeds anything supplied by these older methods [12, 15].

Q–technique was originally introduced by Sir Cyril Burt [2], though its description as Q–technique (or inverted technique[1]) and its popularization among clinicians in the last two or three years has been largely due to Stephenson [23, 24]. This method of correlation and factorial analysis correlates persons instead of tests. Consequently, whereas in the usual R–technique we end with a definition of factors in terms of tests and a test which is a relatively pure measure of some unitary factor, e.g., an intelligence test as a measure of the second order general ability factor, we finish a Q–technique research by pointing to a group of persons as the carriers of a particular factor and perhaps to some one person as supremely exhibiting it. The virtues and defects of this will shortly be discussed.

Criticism of Qs–technique

The controversy between Burt [3] and Stephenson [24] brought out the fact, that Stephenson was using it in two distinct senses, one of which Burt considered illegitimate while the other, he pointed out, yields no results different from those of R–technique. For clarity of future discussion, I shall refer to the second and first of these as Q and Qs — meaning regular Q–technique and *subjective* Q–technique respectively—as I have done in my recent book on factor analysis in the social sciences [23]. Regular Q–technique measures the individuals on objectively scorable, behavioral variables. The results are brought to standard scores and can thus be ranked in order within the individual. For example, individual A may have a standard score of +1.2 on intelligence, –0.3 on sociability and so on. Within that person in-

[1] Burt (3, p. 168) comments on Stephenson's formulation "To talk of inverting the factors is misleading. As I have often pointed out, the theorems required for analyzing correlations between persons are not 'inversions' of those required for the older procedure; they are formally identical with them." Actually, as the reference to rows and columns of the score matrix made below indicates, the calculation is merely on a transpose not an inverse of the R–technique matrix, and we shall henceforth refer to it as the *transposed* spstem.

telligence is then higher than sociability and any person who correlates positively with him is also likely to have intelligence higher than sociability.

As Burt has shown [3] the results of regular Q–technique are mathematically transformable into those of R–technique and vice versa[2]. In short, when we have set out our score sheet it doesn't matter whether we start by correlating the columns (tests) or the rows (persons). Whether R or Q–technique is used becomes a matter merely of convenience and this is largely a question of whether we have a few subjects prepared to spend a long time doing a lot of tests (Q) or many subjects each prepared to spend only a short time, on a few tests (R). The clinician having his patients under lock and key will naturally favor the former; but it is a mistake to suppose that by some statistical magic Q–technique will make the results obtained from a small sample as generalizable as those from a large sample. (Parenthetically, clinicians have shown themselves peculiarly vulnerable to temptations to quick and easy research. The various small sample methods will give answers to questions, but they will not answer quite the same questions as a more thorough research will permit.) By and large, from the same number of subject-hours one can extract the same precision of information by R–technique as by Q–technique.

It is surprising that some trained psychologists have uncritically accepted certain rash claims for special values of Q–technique in the clinical field in spite of the above obvious relationships, pointed out over a decade ago by Burt [3], Thomson [25], Davies [26] and others. To ask that responsible investigators read Burt's classical chapter (3) on this subject is no more than to ask that a person should not invest a great deal in a business before enquiring as to its finances. And it is surely not too much to ask that investigators planning researches in 1950 should have gotten around to reading and reacting to what was written on the subject in 1940.

When the above main relations of R– and Q–techniques have been set out, there remains to state only that the comparison of the two yields four debits for Q–technique as far as most clinical research is concerned, as follows:

(1) It is generally easier to recognize and analyze the nature of a factor from a test than from a person. Inspection of a highly "g" saturated test, say a series of analogies, will readily suggest that relation eduction and correlate eduction are the predominant mental process involved, but inspection of a highly "g" saturated person is another matter, for the intelligence is constantly at the service of different drives.

(2) It is easier to file, preserve or mail.

(3) The first factor (that which would be equivalent to overall size in a factorization of physical measurements) is lost in Q–technique. If two people have the same profile of test scores but at different absolute levels they are identical as far as the eye of the correlation coefficient is concerned: it gives a correlation of unity. Thus the first "general size factor" is overlooked. This might be overcome by using the new "pattern similarity coefficient" [8] except that this has not yet been proven a fit index for factor analytic work.

A statistician will of course point out that a factor is also lost in R–technique relative to Q–technique, because of the usual reciprocal relationship. But the first factor of Q–technique, missing in R–technique, is one in which we are not usually interested. It is the first common factor, indicating the pattern which is common to all people. If the psychologist has mixed rats, cats, apes

[2]It is not necessary to be a statistician to follow all of these arguments, since the issue can be stated also in logical and psychological terms. Regarding Stephenson's claim that Q–technique is *the* method for finding types, which are different from those which would be based on R-technique, Burt points out "To declare that by correlating persons we can discover types that cannot conceivably be discovered by correlating traits is like declaring that there are classes (logical classes) for which no defining predicates (i.e. no trait combinations) could be possibly found." (3, p. 201) From a psychological standpoint, as I have discussed elsewhere [4], a trait is not a test performance but *a set of relations between persons and tests.* Consequently when we factor analyze for source traits it is essentially irrelevant whether we use the columns of tests or the rows of persons: the *scores* used are the same and the whole matrix represents all that is known about the relations of tests and persons.

and people in a common experiment the nature of this common mammalian factor may be of some interest, but usually the clinicians assume that their patients are people and except in extreme instances are not inclined to question it! What they are interested in is how they vary *as people*. The first factor in R–technique, on the other hand, cannot be lost without losing a substantial part of the ability to predict individual differences.[3]

It is sometimes argued that changing raw test scores to standard scores before applying R–technique, to make it comparable to the change to standard scores which is usual before using Q–technique, also loses a factor of information. The contention is that when we bring all tests to the same mean and standard deviation we overlook the information conveyed by the raw scores that our population is at quite different levels in, say, intelligence and sociability. This may happen in a sample, but we are interested in laws applying to a whole universe of human subjects. It is meaningless to say that people are more intelligent than they are sociable, since our standards are gathered from human experience. Or again to make the statement that men are more aggressive than they are neurotic is to say exactly nothing.

(4) It is impossible to rotate the Q–technique factorization to get psychologically meaningful factors by simple structure, at least as long as they remain person-factors, because there is no reason to expect that any factor will absent itself entirely from or be exactly average for many people—indeed the contrary is axiomatic.

Compared with these embarrassing but not insuperable difficulties of Q–technique, which merely put it at a slight discount as a useful instrument, the flaws of Qs technique are totally disabling by any truly scientific standards. Here, as typically used by Stephenson, Rogers and others, the subject is given a list of personality traits and asked to arrange them in their order of "importance" within his personality. By a rank order coefficient this order is then correlated with that of another person (or the same person on a different occasion). It will be observed (1) that this retreats from the scientific standards of behavioral psychology back to the era of introspection in private worlds, and (2) that it is an ipsative measurement [4], i.e. it implies, in, for example, a ranking of interests, that the total interests of one man are as strong as the total interests of another.

In regard to the first the situation could be saved by saying that one is dealing only with the behavior of making marks on a sheet of paper, but this the clinicians who are at present promoting the method decline to do. They take the introspective statements at their face value and assume the subject is giving real data about his personality.[4]

[3] The above was stated more tersely ten years ago by Burt (3, p. 198). "The 'general factor for persons' will describe a characteristic *pattern* of measurements that distinguishes the universe of individuals for whom the factor is 'general' or generic, e.g. in human psychology the human genus." The greater part of Burt's comprehensive survey of Q-technique and his criticisms of Stephenson's misconceptions therein will be found within pages 174-182 of the work cited. The reader should be warned of a possible danger of accidental confusion, through Burt's using Q- and P-technique where I have used Q and Qs technique. Stephenson and others did not adopt Burt's use of P, and I and my students have since found the label P-technique more apt for the entirely new technique of longitudinal analyses of the single person [5, 9, 13, 15, 27].

[4] One must distinguish between using introspective evidence as a basis for hunches that are later to be behaviorally tested, and attempting to use it as part of the architecture of science itself. The writer is not disagreeing with the former: the rough parallelism of actual behavior and self-rating often justifies, with well balanced, honest and insightful subjects, a preliminary run of research at this level to see if results justify the more exact and thought-demanding design in reproducible objectively measurable behavior. But clinicians rarely have such subjects, and the questionnaire type of response is then likely to fluctuate wildly and unpredictably with each chance wind of changing circumstances. (For example, the pre and post therapeutic self ratings may differ merely because of the changing dependence on the therapist.) Moreover, in this field there is no great justification for falling back on introspective evidence since objective tests exist for measuring personality factors before and after therapy, while the structure of the ego and super ego are inferrable by a good clinician from observable behavior.

They also assume that the mystical concept of order of "importance" for personality has an exact meaning and indeed the same meaning for each person. In true Q–technique a common dimension for all persons is indeed created by the standard score, which constitutes a scale within the person of "degree of eccentricity" e.g. his standard score shows he is more *deviant from the average* in intelligence than he is in, say, neuroticism. An alternative suitable term for this scale in objective measures would be "scale of saliency". But the interpretations of *subjective* "importance" will naturally differ for different persons, so that we get *simultaneously variation of the direction of the dimension and of the rank order of the variables upon the dimension.*[5]

Neglect of the Covariation Chart

We thus have, on the one hand, Q–technique, founded on validatable behavior data and yielding methodologically precise results transformable with those of R–technique, and, on the other hand, Qs technique, which rests on a quagmire of unproven—and mostly forever unprovable—assumptions and yields results uncapable of integration with other psychometric measurements or the concepts founded upon them. The distinction is fundamental to an experimenter trained in scientific method, but unfortunately has not been seriously thought upon by many clinicians, who reply that it cannot be wrong because "everyone is doing it" and they are "getting results!"

It would be a thankless public service to ask why the level of research in clinical psychology is what it is and has been during the past fifteen years. But one cannot avoid the diagnosis that the disuse of P–technique and the misuse of Q–technique under discussion here is part of a wider disease in clinical psychology. This mental debility shows itself in such symptoms as an uncritical susceptibility to contagion of fads and fashions; in a dreary, repetitious concentration on such patent medicine tests as the Rorschach and Szondi, when a whole world of new objective tests—at least fifty of proven validity [6, 14] awaits a lifted gaze; and in an evasiveness in regard to facing crucial statistical if not experimental tests of whatever cult of crystal ball gazing is most talked of at the moment. One cannot escape the conclusion that clinical psychology is suffering from the dictatorship of a half-trained intellectual proletariat and that in these circumstances the reply to the above "scientific argument from popularity" is that if everyone is doing it, it must belong to the pre-scientific and inaffective stage of clinical psychology out of which a minority of truly progressive clinicians is trying to rise.

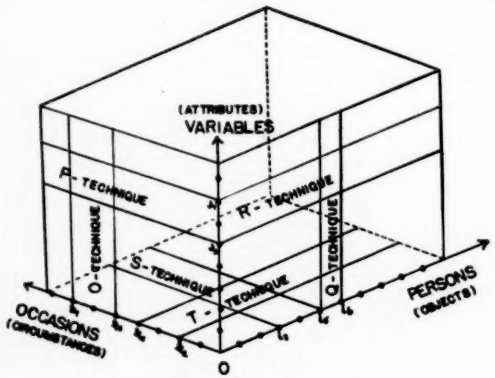
For the confusion of Q and Qs techniques extends also to the inclusion therein of statistical designs which are not Q–technique in any form! A comprehensive analysis of possible correlation designs for research in clinical psychology, personality study and social psychology was published by the present writer in 1946, introducing the concept of the *co-variation chart* [4, p. 97] and has been refined more recently [14a]. Indeed the covariation chart can claim to distinguish beyond reasonable chance of confusion all the *logically possible* covariation combinations possible in any scientific

[5]Burt quotes Stephenson on this "importance" concept and comments as follows (3, p. 186). I have re-written Q as Qs to correspond with the symbolism in this article: "The measurements with which Qs-technique is concerned must be expressed in terms of an entirely new unit, namely *'significance.'* This concept of significance is peculiar to Qs-technique and new to systematic pyschometry . . No better example could be afforded of the difference between Qs-technique and R-technique" (Stephenson (136, p. 233f). One trait is said to be more significant than another if it is "more representative or characteristic of a personality as an indivisible whole (Stephenson (92, p. 293) . . . 'Significance' seems to me merely a matter of weighting."

Unquestionably "significance" is what distinguishes Qs from the interchangeable Q- and R-technique systems, as Stephenson says. It also distinguishes Qs from logic and science. I hope that further thought will confirm Burt's conjecture that it can be reduced to "merely a matter of weighting," but to me Stephenson's phrase "more representative or characteristic of a personality" could mean a hundred things, the only one of which that corresponds to anything in nature would be the "eccentricity" dimension of Q-technique above. But this Stephenson explicitly rejects. If many professional psychologists find the meaning of significance obscure, how much more woolly and unreliable must be the senses in which various patients use it when they are asked to rate themselves in Stephenson's and Rogers' experiments!

area. Therein the R–, Q– and P– techniques which had been formulated as distinct methods about a decade before (though R and Q were much older in actual use) could be seen to fall into place clearly in relation to one another and at the same time *a rich variety of unsuspected and unused additional correlation designs* was revealed.

The construction of this Pandora's box is actually extremely simple, as Fig. 1 shows. One has to remember that one is dealing essentially with three signatures to every psychological act: a person who does it; a "test" or stimulus situation-performance relation which defines it; and an occasion which locates it in time. (Location in space is irrelevant if the stimulus situation is defined.) Any pair of parallel lines drawn in this parallelopiped constitutes a pair of meaningfully correlatable series and therefore a correlation coefficient capable of contributing psychological information. There is no space here to enlarge on examples, but the reader may appreciate fuller definition and discussion in the original presentation[4] or a more recent brief, less fundamental account[11, p. 31] or in Mowrer's recent review[20] or in two recent extensions[13, 14a].



Covariation Chart

A quotation from the original presentation may best sum up what has happened historically in relation to this chart. "Factor analysts have concerned themselves almost exclusively with the plane OP-OV; experimental psychologists have crept gingerly over parts of the plane OP-OO; while clinical psychologists have dealt freely, though without any mathematical method or accuracy, with covariation in the plane OV-OO."[4] Actually there are six primary possibilities of correlation (O, P, Q, R, S, and T) and twelve further possible designs using increments or differences. If one factorizes controlled and independent variables in the new hybrid research designs, combining experiment with factorization, there are still more possibilities.

The Meaning of O-Technique

Among the six primary designs were two, 3a and 3b, styled "Intra-individual Covariation." The first has been used in several pioneer researches[1, 5, 9, 27] and has been described afresh as P–technique above, while the second may be quoted [4, p. 99] as follows: "*Intra-individual correlation* (*and analysis*) *of occasions.* The scores (or rank orders) of a series of tests given to one person are correlated with the same series given on another occasion. The factors would be positions of equilibrium of the total personality with respect to the objective study of multiple personality, in regard to the slighter manifestations of such equilibrium positions in 'normal' persons."

Since this is now coming into practical use it would perhaps be wise to label it, briefly and beyond confusion as *O-technique* or the correlation of *occasions* within the single person.[6] It has recently been used by Rogers to correlate the pre-treatment with the post-treatment personality of a given individual (or the individual's rating of his actual and his ideal self). But the preliminary notices of this research reveal that the clinicians concerned are under the illusion that they are using Q–technique! Q–technique, as defined by Burt, Stephenson, the present writer and anyone else who has actually carried out Q–technique analysis is defined as the "correlation of persons" or "the transposed factor technique" (substituting persons for tests). R technique and its transpose—true Q technique—deal with *several* people and *common* traits whereas P– and O–techniques on the other hand belong to the "statistics of the individual" of which clinical psychology has been so long in need. The blurred focus once more introduced by the clinical touch may perhaps seem excused by the argument that two different persons—the pre-and post-treatment individual—are being compared. But this, apart from its optimism, does not make sense methodologically and the experiment in questions falls into the same pitfall as Qs technique, namely in supposing that two different people can be measured from the introspections of one.

If we are claiming to be responsible investigators, let us at least keep our symbols clear, even though we erect debatable or doubtful experimental designs upon them. By a person we have always meant an individual organism, and test-retest correlations, such as are used in P– and O–techniques, are properly regarded as the study of variations within the same organism. It is not possible to cover up the mistake of considering this a Q–technique design by resort to the casuistry of saying that the pre and post therapy individual, or the individual's conceptions of himself and his ideal self are different persons. They constitute behavior emanating from the same person and in the latter case from the same person at the same moment.

The misuse of O–technique on subjective data also stumbles into one of the pitfalls of Qs–technique mentioned but not stressed above. When each of a series of persons ranks a series of traits, aesthetic choices or interests in the order of significance for himself in the "Q-sort" procedure the assumption is being made by the experimenter that all these individuals stand at the same mean. This can be made an approximation to the truth by careful attention to a stratified sampling of traits, on the basis of the personality sphere universe mentioned above. But in general this careful attention to the R-technique picture has not guided Os-technique and in any case it could only reduce the error systematic to the technique. It is surely no more sound to assume that the same individual has the same strength of total interest on different occasions, in Os-technique, than to assume that the total interests of two people are equal, in Qs technique.

In short, even if we get our symbols clear and get back to the firm ground of distinguishing Q–, P– and O–techniques, the use which Rogers is proposing to make of O–technique produces the same degradation of the method as Stephenson fell into when he substituted Qs–for Q–technique. It constitutes a, seemingly unconscious, regression from behavioristic psychology to introspection, and so to the long nightmare of solipsism from which Hume rescued philosophy and psychology in 1739.

Correct O–Technique and the Self Sentiment

Approached behavioristically, i.e. through *true* measurements, to which there can be more than one witness, as in other sciences, the comparison of pre and post therapy personality structures, or of the ego and the ideal self, would be an obvious and valuable contribution to clinical psychology. If multiple personalities exist within one personality-organism, O–technique will reveal this fact. As I have suggested elsewhere, (11, pp. 645-662, but especially p. 657) if the self sentiment is thor-

[6] When design 2a (*op. cit.*) is discovered by clinicians it will be necessary to call it Ts-technique, because it also correlates occasions but for groups.

oughly analyzed, as to its implications for dynamic subsidiations, anyone with more than a layman's grasp of the self concept should be able to set up *behavioral* measures from which the nature and relationships of the actual and ideal self, or two "multiple personalities" can be inferred. Indeed, a statistical analysis of a set of fifty objectively measured attitudes (using memory effects, P.G.R. measures, etc.) on two hundred subjects has already revealed a pattern of reinforcement of ordinary interests which can most simply be interpreted as arising from the action of the self-sentiment [12] and this analysis provides behavioral evidence of the existence of the self sentiment.

When such comparisons of personality structures are made, even with sound objective measures, there is a need to choose the variables with regard to meaningful hypotheses of personality structure and to the statistical nature of the Q–, P– and O–techniques. In the clinical researches now in imminent danger of becoming fashionable, and especially in the Qs–techniques of Stephenson, this has not been done. Stephenson uses Qs-technique to divide humanity into types. Each individual lists a set of personality variables "in order of importance" for his personality, and individuals whose orders correlate highly are said to be of the same type. But there is no guarantee that different aspects of personality are equality represented. If 90% of the variables deal with esthetic tastes two individuals whose esthetic training happens to coincide will correlate as highly similar though in fact they differ in many other and more important aspects of personality more than do people indicated as less similar! Similarly in O–technique if we want to say that an individual is more like himself (his model self) in situation-occasions a, d and e than in a and c, the series of variables upon which we re-measure him on all these occasions must be properly distributed.

In fact, a logical analysis will show that we are thrown back on the personality sphere [4] concept (or its equivalent) in R–technique as our only guide. That is to say, in O–technique, or in Q–technique used to define types, *it is necessary to use as variables the personality factors which have already been isolated as primary in R–technique.* For example, we should ask how much the individual varies from day to day in emotional stability (C-factor), in schizoid withdrawal (H-factor) and so on through the various functional unities, instead of blindly discharging a shot gun of variables a majority of which may fall only in one factor. Incidentally such an eye to distribution is desirable also in P–technique; but that method does not stand or fall thereby as do Q– and O–techniques when used with the theoretical objectives of these clinicians.

R–Technique Basic to P, Q and O

R-technique thus remains the *fons et origo* of Q–, P– and O–techniques. Some of the recent writing on the clinical importance of the "statistics of the individual" would seem prone to forget this and there have indeed been some egregious examples of the fact that "men to eke the praise of new loves (will) lie about the old." These new techniques can only be wisely applied if we recognize their true relations to older factor analytic methods and indeed to the whole of statistics, as has recently been pointed out in regard to various research methods by Grant [18]. In the first place, since the factors of R–technique itself have as yet been so negligibly explored for their clinical implications it would probably be sound research strategy to attempt nothing more than reconnoitering studies with these new methods until the R-technique findings are consolidated. In the second it is necessary to keep in mind that the P–technique factors are the *unique* trait patterns which should center upon and disperse about the community-anchored R–technique trait patterns. Unless the R–technique patterns have been well and truly stabilized the present work with P– and O–techniques has no foundation with which to relate itself, and as far as the growth of knowledge is concerned will be like the seed that fell on stony ground.

Actually, although O–technique has been little guided in this way, four of the five[7] P–technique studies so far completed [1, 5, 9, 15, 27] have deliberately been planned in the light of R–technique findings. In the first demonstration (5) in which a woman was measured on 35 objective tests, 30 behavior ratings and a number of questionnaire self-ratings (not used as essential evidence), great care was taken to represent the then known factors in R–technique, including the behaving rating factors A through N, some "multiphasic" questionnaire factors and what few objective test factors were then known. Correlations among these variables were based on re-examination of the subject every day for nine weeks, a length of series which justified extracting about half a dozen factors. By way of a test of the factorization the objective tests and the ratings were separately factored and the factors then cross-related. A beautifully clear identity then manifested itself, in which three rating factors neatly coincided with three of the four test factors. One of these was a diurnal fatigue factor naturally not found in R–technique, but two others, both in rated behavior and in objective measures, coincided with two of the best known R–technique factors—C, Emotional Stability and F, Surgency. The especial contribution of the P–technique study was in showing the peculiar modes of manifestation of these functional unities in this particular personality, i.e. the unique trait structure.

P–Technique and Individual Dynamics

In regard to clinical investigation—whether it be of general principles or the routine diagnosis of cases—it would seem that P–technique gives promise of being far more valuable than any other statistical tool. Hitherto experimental-statistical devices have been perceived by the clinician as more exact but less capable of coming to grips with his real problems. With the invention of P–technique the clinician has a method which is at once exact and global. By its ability to comprehend in a single calculation all the dynamic manifestations of the personality and to reveal their structural relations and connections, it makes free association and uncontrolled "projective" techniques obsolete—except in so far as the former may be needed, as a therapeutic aid, to provide the patient with insight. Even so, the therapist can aid that insight more confidently if he already knows from a P–technique analysis the quantitative roles of various drives and personality structures in the symptom formation.

The only illustration of P–technique yet presented on a clinical case [9] perhaps did not sufficiently concentrate on dynamic material to give as complete a demonstration of its powers in this direction as would be necessary to alert the average clinician to its possibilities. The study in question had to be designed partly to solve the logically prior issue as to whether R– and P–techniques would yield the same unitary trait patterns, as theory would lead one to expect. This expectation it confirmed, but it also gave sufficient indication of power to reveal dynamic connections, e.g. the correlation of stomach ulcer symptoms with suppression of anxiety, as to leave no doubt of its clinical practicability.

The logic of the use of P–technique in clinical psychology is actually very simple. It is assumed that the drive strength of the primary drives will vary from day to day or week to week with (a) physiological conditions, at least in such drives as hunger or sex, and (b) with changes of stimulation in the environment. Those behavioral

[7] I have included Baldwin's (1) study here, though as I have pointed out earlier not all intra-individual correlations can be considered P-technique, when the latter is properly defined. I believe it would avoid confusion, especially over different investigators getting different results, if we reserved the terms R-, Q-, P- and O-techniques, not for the types of correlation, but for factor analysis erected upon the types of correlation. Thus the correlation of tests on series of persons and the inspection of these correlations for clusters does not constitute R-technique. Nor does the correlation of variables observed longitudinally in the individual constitute P-technique, unless a factor analysis ensues which is rotated for simple structure, admits correlated source traits and sets up other conditions of psychological reality in the factors. Unless this is born in mind it is possible to give the impression of contradiction in the findings of different workers with "the same technique", as Stephenson does according to Mensh (19, p. 481).

variables which are expressions of the same drive will thus show positive correlation in day to day variation. They will also show some tendency to negative correlation when they represent mutually substitutable behavior, i.e. when they are dynamic equivalents. Simple examination of all the correlation coefficients among the various dynamic expressions and symptoms is not capable of revealing such complex connections. But it has been shown [1, 5, 9, 15, 27] that factor analysis can do so, and that we then obtain factors corresponding to the drives and to the controlling ego and superego structures, the loadings of which in particular acts, dream expressions and symptoms is a quantitative statement of their roles, permitting a calculus of psychotherapy previously impossible.

Clinical Improvements in P–Technique

In the light of these considerations, Rosenzweig's statement that factor analysis is "not a method of directly appraising or investigating a personality" appears [22] as one more instance of clinician misleading clinician, or at least of bolstering fellow clinicians in a complacency with their present methods. The fact is that P–technique is at present the *only* method which will do with exactness what the clinician is attempting to do by rough, common-sense methods such as free association and observation of change of symptoms with the course of therapy. And R–technique is the only method capable of deciding whether there is any functional unity in such concepts as "general neuroticism," "ego strength" or "free anxiety" which the clinician freely uses.

P–technique analysis, if polished and simplified to clinical practicability, could thus well become the standard diagnostic method for "appraising or investigating a personality." The slight improvements which cry out to be made to complete its clinical applicability are: (1) It would seem desirable to correlate in with the patient's dynamic expressions certain presses or stimuli in his daily environment (including measures on the psychotherapist) which are suspected of being important in his adjustment. (2) Where we suspect a causal lag between influence and response, it would be desirable to experiment with "staggered correlations," i.e. matching the series for the given two variables by putting Monday's measure of one with Tuesday's measure of the other, and so on, before correlating. (3) It would be desirable to include in every P–technique factorization (now that some constancy has been shown in P–technique factors and in their relation to R–technique factors) a standard set of reference variables constituted by those variables which have been found to be the mostly highly-loaded measures of the basic drives, of ego and superego structures, etc. This last device, as improved knowledge of structure and of best-loaded variables is gained, will eventually obviate the necessity for carrying out an actual factorization when using P–technique in routine clinical diagnosis. For, as a statistician will realize, it will suffice simply to correlate each of the other variables with these standard "factor representatives," then to expand the latter to unit length and finally to make whatever slight rotational adjustments are necessary. Since one can only conclude that the failure of the average clinician to avail himself of P–technique (and also of properly constituted Q–technique) is due to an aversion to skilled calculation and the hard work of acquiring new techniques such a simplifying step might place P–technique within his reach.

Summary

(1) The necessarily limited role of experiment in clinical research throws the onus of proof upon statistical methods and of these the only "global" method capable of dealing with the wholes familiar to the clinician is factor analysis, particularly the R–, Q–, P– and O–techniques.

(2) The first two deal with groups, the results of one being transformable into the other, and each having certain specific conveniences in research. But P– and O–deal with, and are the ideal methods for, the single case. O–technique correlates the personality sphere of the individual on one occasion with himself on another occasion.

It is essentially the correlation of *occasions*—hence O–technique. With variation of circumstances through a whole series of occasions it is possible to group the occasions according to the factors they manifest and share. It is thus the required method for the investigation of the sociologist's theories about the effects of roles upon personality and for detecting in clinical investigation the circumstances which emphasize different aspects of the self.

(3) At present clinicians seem to know these methods only through two degenerate forms, Qs and Os techniques, which abandon behavioristic standards of evidence and make various invalid assumptions, preventing the use of the results in the same frame of reference as that of general psychology based on the study of behavior.

(4) P–technique, which factorizes a set of variables within a single individual, after testing repeatedly to create a longitudinal series, is the method which is intrinsically most relevant to clinical work and which promises to bring both investigation of principles and diagnosis of the individual to a new level of scientific accuracy. Its most obvious use is in establishing dynamic structure and in tracing symptoms to ergic origins. Put P– and O–techniques are most effectively employed only after the common structure of personality has been established by R–technique, thereby providing a proper perspective for sampling personality sphere variables.

(5) The above are only four out of some eighteen or more logically possible ways of utilizing covariation observations to throw light on personality structure. These possibilities have been precisely described in the *covariation chart*, from which clinicians might obtain methods apt to throw light on a wide range of resistant problems. Sources of reading are indicated which might remedy both the disuse of this range of methods and the misuse of the few that have been tried, but this suggestion assumes that responsible clinical researchers will be less interested in quick and easy than in truly effective methods, capable of advancing architectonic research.

Bibliography

1. Baldwin, A. L. Personality structure analyses: A statistical method for investigating the single person. *J. Abn. Soc. Psychol.*, 1942, 37, 163-183.
2. Burt, C. L., Gaw, F., Spielman, W., and Smith, M. *A Study in Vocational Guidance.* London: H. M. S. Office, 1926.
3. Burt, C. L. *The Factors of the Mind.* London: University of London Press, 1940.
4. Cattell, R. B. *The Description and Measurement of Personality.* New York: World Book Company, 1946.
5. Cattell, R. B., Cattell, A. K. S. and Rhymer, R. M. P-technique demonstrated in determining psycho-physiological source traits in a normal individual. *Psychometrika*, 1947, 12, 267-288.
6. Cattell, R. B. Primary personality factors in the realm of objective tests. *J. Person.*, 1948, 16, 459-487.
7. Cattell, R. B. *A Guide to Mental Testing.* London: University of London Press, 1948.
8. Cattell, R. B. rp and other coefficients of pattern similarity. *Psychometrika*, 1949, 14; 279-298.
9. Cattell, R. B. and Luborsky, L. P-technique demonstrated as a new clinical method for determining personality and symptom structure. *J. Gen. Psychol.*, 1950, 42, 3-24.
10. Cattell, R. B. and Saunders, D. R. Inter-relation and matching of personality factors from behavior rating, questionnaire and objective test data. *J. Soc. Psychol.*, 1950, 31, 243-260.
11. Cattell, R. B. *Personality: A Systematic Theoretical and Factual Study.* New York: McGraw Hill Book Company, 1950.
12. Cattell, R. B. The discovery of ergic structure in man in terms of common attitudes. *J. Abn. Soc. Psychol.*, 1950, 45, 598-618.
13. Cattell, R. B. *Factor Analysis for Psychologists.* New York: Harper Brothers Company, 1950.
14. Cattell, R. B. A factorization of newly designed objective tests of personality source traits. *Brit. J. Psychol.*, Statist. Section, 1951.

14a. Cattell, R. B. The three basic factor analytic designs—P–, R– and T– techniques—and their derivatives. (In Press). *Psychometrika.*

15. Cross, K. P. A P-technique analysis of objectively measured attitudes in regard to a theory of ergic structure. Urbana: University of Illinois Library, 1951.
16. Davies, M. 1939. The general factor in correlations between persons. *Brit. J. Psychol.*, 1939, 29, 404-421.

17. EYSENCK, H. J. *Dimensions of Personality.* London: Kegan Paul, 1947.
18. GRANT, D. Statistical theory and research design. In C. P. Stone (Ed.), Annual Review of Psychology, 1950. 277-296.
19. MENSH, I. N. Statistical techniques in present day psychodiagnostics. *Psychol. Bull.*, 1950, 47, 476-492.
20. MOWRER, O. H. *Q Technique—Description and Critique in Theory and Research in Psychotherapy.* New York: Ronald Press, 1951.
21. ROGERS, C. Personal Communication. Also M. Hartler's mimeographed form for clinicians "Q-technique: its methodology and application".
22. ROSENZWEIG, S. 1948. Available methods for studying personality. *J. Psychol.*, 1948, 28, 345-368.
23. STEPHENSON, W. The foundations of psychometry: Four factor systems. *Psychometrika*, 1936. 1, 195-210.
24. STEPHENSON, W. A statistical approach to typology: The study of trait universes. *J. Clin. Psychol.*, 1950. 6, 26-38.
25. THOMSON, G. H. On complete families of correlation coefficients. *Brit. J. Psychol.*, 19365. 26, 63-92.
26. THURSTONE, L. L. 1947. *Multiple Factor Analysis.* Chicago: University of Chicago Press, 1947.
27. WILLIAMS, H. VER MEER. A P–technique study of personality factors in the psychosomatic areas. Urbana: University of Illinois Library, 1949.
28. ZUBIN, J. Introduction: Symposium on statistics for the clinician. *J. Clin. Psychol.*, 1950, 6, 1-6.

RESEARCH IN CLINICAL PSYCHOLOGY: 1950

WILLIAM SCHOFIELD

University of Minnesota

In reporting on a review of a reasonably representative sample of the research in clinical psychology published in 1949, the writer was encouraged to comment on several hopeful signs.[30] These signs included not only an increase in the quantity of research appearing in the major journals, but evidence that after a considerable period of concern for professional status and standards, our discipline was at long last turning a critical eye to the examination of techniques which had gained widespread acceptance predominantly on the basis of authority. A similar review of the research published in 1950 permits the writer to feel that he was not premature in his earlier expression of hopefulness.

To facilitate comparisons the same journals included in last year's survey were used for purposes of this review: *Journal of Clinical Psychology, Journal of Consulting Psychology, Journal of Abnormal and Social Psychology, Journal of Applied Psychology, Journal of General Psychology, and Journal of Psychology.* These six journals yielded a total of 148 clinical researches as defined by the same screening criterion previously applied, i.e., they were studies "involving the systematic investigation of a specifically described group of subjects and the derivation of normative or comparative data from psychometrics, case histories, or therapeutic interviews, or analysis of administration, scoring, and interpretation of a given instrument." These 148 studies represent an increment of over seventy per cent of the number (86) found in the same journals in the preceding year.

A brief abstract was prepared for each of the studies and these abstracts were later used in sorting the studies into relatively distinct areas of research. Again to facilitate comparison, the basic set of categories developed for the previous reviews was applied. Table 1 reports the distribution of the 148 studies among the various categories with comparable data for the 86 studies reviewed in 1949. It is interesting to note that while certain categories of research have held a prominent position with

Table 1. DISTRIBUTION OF 148 RESEARCH STUDIES REPORTED IN SIX SELECTED JOURNALS IN 1950, BY AREA OF RESEARCH REPRESENTED, WITH COMPARATIVE DATA FOR 1949.

	Area	No. of Studies	Per cent of Total	Per cent of Total 1949	Rank 1949
1.	Validity (projective)	29	19.6	17.4	1
2.	Normative study (projective)	16	10.8	2.3	12
3.	Objective evaluation of therapy	14	9.4	7.1	4
4.	Normative study (intelligence)	13	8.9	4.7	8
5.	Intertest relationships	13	8.9	5.8	6
6.	Validity (diagnostic patterns, W-B)	11	7.4	4.7	7
7.	Normative study (personality)	7	4.7	9.3	3
8.	Test standardization	7	4.7	2.3	16
9.	Validity W-B Deterioration Index	5	3.4	2.3	15
10.	New tests (projective)	5	3.4	2.3	17
11.	Normative (structured personality tests)	4	2.7	3.5	10
12.	Abbreviated intelligence tests	3	2.0	5.8	5
13.	Validity (prognostic indicators)	3	2.0	2.3	14
14.	Validity (structured personality tests)	2	1.4	2.3	11
15.	New tests (intelliegnce)	2	1.4	2.3	18
16.	Detection of malingering	1	0.6	2.8	19
17.	Analysis of recorded interviews	1	0.6	9.3	2
18.	Differential diagnosis of feebleminded	0	0.0	3.5	9
19.	Validity of psychiatric diagnosis	0	0.0	2.3	13
20.	Miscellaneous	12	8.1	8.2	20
		148	100.0	100.0	

respect to frequency of studies others have dropped considerably. Validity studies of projective techniques continue to hold first place, and the objective evaluation of therapy remains within the first five categories in order of frequency. On the other hand, the analysis of recorded interviews drops from second to seventeenth place and accounts for less than 1% of the reviewed studies. Likewise, abbreviated intelligence tests drop from fifth to twelfth place. Normative studies of projective techniques are the second most frequent type of research, jumping from twelfth to second place and accounting for over 10% of the reviewed studies. One-third of all the studies are concerned directly with projective techniques.

The first category, validity study of projective techniques, warrants special comment. The 29 studies included fifteen studies of the Rorschach, three studies of the TAT, four Szondi studies, three studies of the Bender-Gestalt, two studies of figure drawing, one study of the Rosenzweig P-F Study, and one study of the Hanfmann-Kasanin Concept Formation Test. Of the fifteen Rorschach studies, ten yielded negative findings, two yielded results of borderline significance, and the remaining three were interpreted as supporting the contentions that the Rorschach measures "basic personality structure" rather than superficial situational variables, that depression is reflected in Rorschach protocols by emphasis on pure form, by "form with textural elaboration (Fc)," and by lack of M and Fm, and that the Rorschach can be effectively added to the ACE in predicting college grades.[9, 6, 28] In the last study, using an optimal combination of weighted Rorschach scores selected from the Harrower multiple choice technique, plus the ACE, a multiple correlation of +.62 with college grades was obtained; this is compared with a zero-order correlation of ACE with grades of +.56. The authors conclude that the addition of the Rorschach makes for "appreciably better" prediction of college grades. They do not come to grips with the question of whether the difference between the coefficients of +.56 and +.62 is statistically reliable or, what is more important with consideration for the time and effort involved in Rorschach testing, practically significant.

Wittenborn's studies again deserve notice. Using a factor analytic approach to study the mechanics of the Rorschach, Wittenborn is concerned with whether the commonly scored and interpreted Rorschach variables are interrelated as they should be if these scoring and interpretive approaches were valid.[36] He concludes that for the most part the Rorschach behaves like a test, i.e., like an instrument capable of reflecting differences among individuals and revealing something of the inner states or reaction tendencies of individuals. Exactly what the Rorschach measures and how is not the immediate focus of Wittenborn's attack. Nevertheless, his findings are such as to discourage the facile assumption that the answers to these problems of external validity have already been given. Thus, his findings do not support the generally accepted practice of distinguishing between CF and FC responses nor do they justify the usual scoring and interpretive distinctions between vista, texture, and color responses. Wittenborn finds all of the latter three categories plus whole responses to constitute a single factor which he tentatively identifies as "perceptual control." In another study using chi square technique to investigate the internal consistency of the Rorschach scoring categories, he finds justification for the orthodox practice of totaling color responses and assigning them an interpretation different from that given to human movement responses; also he finds justification for the practice of assigning distinctive interpretation to the various categories of the location variable.[37] However, his findings lead him to make the following very important reservation: "The present data are in no way a justification for the common practice of ascribing important personality differences between individuals as a result of small differences in the frequency with which color responses, movement responses, whole responses, or large detail responses appear in the Rorschach protocol of the individual."

Factor analysis was applied in another study by Hughes who was interested in determining the validity of certain of the more common diagnostic signs read from

the Rorschach.[21] From a careful review of the literature, Hughes culled 22 signs claimed to differentiate intracranial pathology, psychoneurosis, and schizophrenia. He determined the presence of these signs in the protocols of 100 cases nearly equally distributed among these three diagnostic categories. A factor analysis of the matrix of tetrachoric intercorrelations of the signs produced eight factors. One of these factors was heavily loaded with "organic" signs; when the factor loadings on this factor were assigned weightings, the resulting scores permitted identification of 82% of organic cases at the cost of misidentification of only 1% of non-organics. By contrast Piotrowski's ten signs of intracranial pathology correctly identified only 20% of the organics. Hughes shares Wittenborn's conviction as to the value of factor analysis in helping to establish an experimentally rigorous footing for the Rorschach technique as a preliminary to validity study of its behavioral referents. Like Wittenborn he permits optimism for the potential of the Rorschach but not complacency over its established worth: "The present study while confirming the relationship between objective Rorschach signs and clinical diagnosis reveals that signs now in use do not distinguish with any degree of accuracy between psychoneurosis and schizophrenia."

In other Rorschach studies, *negative* findings were reported for the validity of the variables commonly interpreted as related to the shyness-gregariousness axis of personality, for the signs of homosexuality, for the usual interpretation of the relationship of color sensitivity to productivity and reaction time, for the Muench "adjustment" signs, for the human movement and other movement responses as predictive of intelligence, and for the differential "affective value" of the ten cards.[20, 12, 10, 19, 33, 24] Carp and Shavzin in a study of the "fakability" of the Rorschach found individuals who, under instructions to produce "good" or "bad" impressions, could manipulate their responses and vary their personality picture as seen in the Rorschach.[8]

The two studies of figure drawing which occur in the first category of Table 1 are of particular interest in that they represent respectively a validity study of certain diagnostic signs derived from an essentially intuitive approach to this projective technique and, on the other hand, a demonstration of a truly psychometric approach to the same type of material. In the former, Fisher and Fisher studied the reliability with which trained persons could rate for the presence of six of Machover's paranoid signs, and the accuracy with which such signs permitted the detection of paranoid trends in diagnosed schizophrenics of paranoid type.[15] In only 18% of the 32 cases was there agreement by all three experienced raters as to the presence of a paranoid component. In another part of this study, ratings of the facial expression and stance of drawings were obtained from psychologists trained in figure drawing analysis, from untrained psychiatrists, and from stenographers. The trained psychologists showed no greater agreement among themselves than they did with the untrained psychiatrists or with the stenographers. The writers conclude that their "total results suggest that it is precarious to accept most of the current assumptions regarding figure drawing analysis (particularly as expounded by Machover) without the confirmation of further research."

In the second study on figure drawing, Albee and Hamlin demonstrate that with the application of simple psychological measurements to the projective material of this technique, valid as well as clinically useful data are derivable.[2] In a previous study, they used the paired-comparison technique with 15 psychologist-judges in order to derive a criterion scale of ten drawings covering the "adjustment" continuum. In this study, this 10-step scale was used to evaluate the figure drawings of three groups of subjects: out-patient schizophrenics, out-patient anxiety neurotics, and VA dental clinic patients. The drawings were presented in random order to four judges who were required to assign each drawing to a point on the scale. The reliability of the judges, considered as two pairs of two judges each, was +.89. Reliable differences were obtained between the mean scale "adjustment" values of the normal and neurotic and the neurotic and schizophrenic groups respectively. These

two studies very well characterize current research on the validity of projective techniques. We are seeing the consistent piling up of evidence against the validity of those instruments and modes of interpretation which have been fostered by clinicians seeking to reap a rich harvest without tilling the soil, content to believe that if the seed is present the plant is inevitably forthcoming. On the other hand, there is a growing number of researches in which workers willing to engage in unglamorous tilling activities demonstrate convincingly that the projective seed can bear edible fruit.

This state of affairs should make for a generally sanguine outlook. It is obvious, however, that the issues are far from clear cut for many of the participants. Thus, Guertin opens a paper on the Szondi "test" with the following arresting statement: "Obviously, the fate of a new projective technique should not be determined by the practicality of a direct validation study."[16] The full import of this statement involving as it does questions of the validity of criteria, the reliability and objectivity of scoring systems, the adequacy of standardization groups, etc., is acceptable. The initial reading, nevertheless, leaves wonder as to the "obviousness" that one should be solicitous for the "fates" of projective techniques independently of their empirically established worth. In his paper, Guertin establishes that when the Szondi records of state hospital employees are scored by the Deri technique the resulting profiles are reliably different from those generated from a random number table. However, in a separate study, he demonstrates that simple visual memory could more economically account for test-retest behavior on the Szondi than the hypothetical need system.[17] His first paper establishes that the Szondi behavior of normals is not a chance phenomenon; his second paper that what aspects of personality the Szondi "measures" and how it does so are not parsimoniously accounted for by the Szondi rationale. The disturbing thing is that instead of the burden of proof resting with Szondi (or Deri), it falls to the already overburdened clinician who finds it *unfashionable to disbelieve*, on the simple basis of insufficient and unconvincing evidence, but is compelled to accept as valid that which has not been disproven!

A completely hopeful outlook for our discipline, in view of the aggressive research now appearing and the gradual displacement of fiction by fact, is further clouded by the continued appearance of obsequities of which the following are typical: "The Rorschach is the clinical psychologist's most important diagnostic technique;" "The Rorschach ink blot test is widely accepted as one of the most acute measures of personality structure available." These incantations seem a strange phenomenon in view of the actual validity status of this technique. They necessarily reflect the acceptance of prestige statements in lieu of evidence and the misinterpretation of ubiquity as synonymous with verified utility. With regard to projective techniques in general, clinical psychology has been on the selling end of a "pig in a poke" bargain. Very belatedly our discipline is trying to make good on the extensive credits that have been extended to it by sister sciences. It should be as much to their discomfort as to our own that they have invested heavily on the basis of persuasive advertising without demanding convincing samples of the real product. The extensive efforts now underway to separate what has been merely claimed from what has been sufficiently demonstrated and to test claimed relationships should not require introductory testimonials when they are published.

Among studies in the second category, normative studies of projective techniques, several papers warrant special mention. Beck, Rabin, et al, present a preliminary paper in what is to be a series concerned with the investigation of schizophrenia by means of the Rorschach method.[3] In this paper they report on the distributions of an extensive list of the major Rorschach variables as they found them in a well described, carefully selected but small ($N = 157$) sample of normal adults of both sexes. These data constitute a very valuable extension and correction of Beck's previously published norms.

Weider and Noller provide data on the relationship of socio-economic level to awareness of sex in the figure drawings of children in the first of what is to be a series

of objective studies of this medium of personality expression.[35] In contrast to the boys, their girl subjects showed a reliably greater tendency to draw figures of their own sex under instruction to "draw a figure." Eron, Terry, and Callahan developed empirical rating scales to cover the range of emotional tone elicited from the Murray TAT cards administered to 50 male and 50 female college students.[11] When it was established that reasonably high reliabilities among judges could be obtained with the application of these scales to new protocols, they were used to investigate the relative stimulus value of each of the TAT cards as elicitors of emotional tone.

In an interesting study of intrafamilial personality configurations, Morris and Nichols obtained Rorschach protocols on both parents of a sample of 46 children diagnosed as "primary behavior disorder," as well as on the children.[27] Twenty variables covering intellectual and emotional attributes and "over-all pathology" were selected for rating from the Rorschach protocols. On fifteen of the variables significant homogeneity among all three family members was demonstrated. The authors explain the finding that the children were rated significantly higher on over-all pathology than either parent as due to the "cumulative reinforcement effect" on the children of pathological traits present in both parents. This hypothesis is not supported by the actual data on the relative "goodness" of the children and their parents on the several traits rated from the Rorschach.

Of the fourteen studies in Category #3, five are reports of attempts to get objective measures of the effects of nondirective therapy. All five report generally positive outcomes for client-centered approaches when measured in terms of such criteria as favorable evaluation by the counselees, improved Bernreuter scores, smaller drop-out rate from college, and significant increases in reading ability. The last constitutes a considerably more impressive though indirect criterion than has generally been applied in evaluation of nondirective treatment.[4, 5] In one of the very few studies to be published to date on the effects of prefrontal lobotomy on Wechsler-Bellevue performance, McCullough reports data on ten patients who were tested one week pre-operatively and were retested in the second and again in the eighth post-operative weeks.[25] The Vocabulary, Information, and Similarities subtests were most resistant to the effects of the operation. The mean scores on the Digit Span and Picture Arrangement subtests showed reliable increases. Studies by Kaufman[22] and Carp[7] of the degree to which the MMPI measures the clinical effects of psychotherapy and insulin shock respectively are of particular interest because of the close manner in which they parallel the earlier findings of comparable studies.[31, 18]

The first three categories of Table 1 account for 40% of all the tabulated studies and the next three categories cover an additional 25%. The contents of Category #6, concerned with the validity of Wechsler-Bellevue diagnostic patterns, deserve review. Of the eleven studies in this category, eight report negative findings. In the 1949 review, four out of four studies gave negative findings. It seems appropriate to ask how many negative studies must be accumulated to convincingly refute the heuristic claims of Wechsler and Rapaport as to the usefulness of W-B patterns for psychiatric diagnosis.[34, 29]

It is not possible to give individual mention to all of the many excellent studies included in this survey. It is desirable, however, to call attention to particular studies which combine clinical ingenuity with lucid experimental design in testing basic hypotheses and measuring the extent of possible practical relationships which are suggested by the hypotheses. Albee's paper on the patterns of aggression in mental patients and the relationship of such patterns to prognosis is such a study.[1]

Finally, mention must be made of a subgroup of papers falling in the miscellaneous category. These are studies dealing with the concepts of psychologists and psychiatrists concerning the nature of therapy and therapists. At a time when psychotherapy is one of the most active issues in our field, it is encouraging to find a number of researches which demonstrate that some of the differences among "schools" are more apparent than real. Seeman found 85% agreement among 70

"expert" psychotherapists as to the desirability that therapists-in-training have access to personal therapy.[32] Mensh and Watson found a high degree of agreement among experienced therapists of diverse orientation as to desirable treatment goals.[26] Fiedler demonstrated that "expert" representatives of the orthodox analytic, nondirective, and Adlerian schools create therapeutic relationships which are very similar as measured against an ideal therapeutic relationship and, furthermore, that their therapeutic relationships were more alike than those created by neophyte therapists of the three schools.[13, 14] These findings suggest that the nature and techniques of effective psychotherapy are probably best to be studied by attention to those aspects of "expertness" which appear to be essentially independent of theoretical orientation.

Bibliography

1. Albee, George W. Patterns of aggression in psychopathology. *J. consult. Psychol.*, 1950, 14, 465-468.
2. Albee, George W. and Hamlin, Roy M. Judgment of adjustment from drawings: The applicability of rating scale methods. *J. clin. Psychol.*, 1950, 6, 363-369.
3. Beck, Samuel J., Rabin, Albert I., Thiesen, Warren G., Molish, Herman, and Thetford, William N. The normal personality as projected in the Rorschach test. *J. Psychol.*, 1950, 30, 241-298.
4. Bills, Robert E. Nondirective play therapy with retarded readers. *J. consult. Psychol.*, 1950, 14, 140-149.
5. Bills, Robert E. Play therapy with well-adjusted retarded readers. *J. consult. Psychol.*, 1950, 14, 246-249.
6. Blake, Robert R. and Wilson, Glen P., Jr. Perceptual selectivity in Rorschach determinants as a function of depressive tendencies. *J. abn. soc. Psychol.*, 1950, 45, 459-472.
7. Carp, Abraham L. MMPI performance and insulin shock therapy. *J. abn. soc. Psychol.*, 1950, 45, 721-726.
8. Carp, Abraham L. and Shavzin, Arthur R. The susceptibility to falsification of the Rorschach psychodiagnostic technique. *J. consult. Psychol.*, 1950, 14, 230-233.
9. Counts, Robert M. and Mensh, Ivan N. Personality characteristics in hypnotically-induced hostility. *J. clin. Psychol.*, 1950, 6, 325-330.
10. Dubrovner, Raphael J., Von Lackum, William J. and Jost, Hudson. A study of the effect of color on productivity and reaction time in the Rorschach test. *J. clin. Psychol.*, 1950, 6, 331-336.
11. Eron, Leonard D., Terry, Dorothy, and Callahan, Robert. The use of rating scales for emotional tone of TAT stories. *J. consult. Psychol.*, 1950, 14, 473-478.
12. Fein, Leah. Rorschach signs of homosexuality in male college students. *J. clin. Psychol.*, 1950, 6, 248-253.
13. Fiedler, Fred E. A comparison of therapeutic relationships in psycho-analytic, non-directive, and Adlerian therapy. *J. consult. Psychol.*, 1950, 14, 436-445.
14. Fiedler, Fred E. The concept of an ideal therapeutic relationship. *J. consult. Psychol.*, 1950, 14, 239-245.
15. Fisher, Seymour and Fisher, Rhoda. Test of certain assumptions regarding figure drawing analysis. *J. abn. soc. Psychol.*, 1950, 45, 727-732.
16. Guertin, Wilson H. A consideration of factor loadings on the Szondi test. *J. clin. Psychol.*, 1950, 6, 262-266.
17. Guertin, Wilson H. A test of a bsaic assumption of the Szondi. *J. consult. Psychol.*, 1950, 14, 404-407.
18. Hales, William M. and Simon, Werner. MMPI patterns before and after insulin shock therapy. *Am. J. Psychiat.*, 1950, 105, 254-258.
19. Hamlin, Roy M., Albee, George W., and Leland, Earl M. Objective Rorschach "signs" for groups of normal, maladjusted, and neuropsychiatric subjects. *J. consult. Psychol.*, 1950, 14, 276-282.
20. Holtzman, Wayne H. Validation studies of the Rorschach test: Shyness and gregariousness in the normal superior adult. *J. clin. Psychol.*, 1950, 6, 343-351.
21. Hughes, Robert M. A factor analysis of Rorschach diagnostic signs. *J. gen. Psychol.*, 1950, 43, 85-103.
22. Kaufmann, Peter. Changes in the Minnesota Multiphasic Personality Inventory as a function of psychiatric therapy. *J. consult. Psychol.*, 1950, 14, 458-464.
23. Kitay, Julian I. The Bender Gestalt test as a projective technique. *J. clin. Psychol.*, 1950, 6, 170-174.

24. LEVY, JEANNE R. Changes in the galvanic skin responses accompanying the Rorschach test. *J. consult. Psychol.*, 1950, 14, 128-134.
25. MCCULLOUGH, MILTON W. Wechsler-Bellevue changes following prefrontal lobotomy. *J. clin. Psychol.*, 1950, 6, 270-273.
26. MENSH, IVAN and WATSON, ROBERT I. Psychiatric opinions on personality factors in psychotherapy. *J. clin. Psychol.*, 1950, 6, 237-242.
27. MORRIS, WOODROW W. and NICHOLS, ALMA L. Intrafamilial personality configurations among children with primary behavior disorders and their parents: A Rorschach investigation. *J. clin. Psychol.*, 1950, 6, 309-319.
28. OSBORNE, R. TRAVIS, SANDERS, WILMA B. and GREENE, JAMES E. The prediction of academic success by means of "weighted" Harrower-Rorschach responses. *J. clin. Psychol.*, 1950, 6, 253-258.
29. RAPAPORT, DAVID. *Diagnostic Psychological Testing*, Vol. I. Yearbook Publishers, Inc., Chicago, 1945.
30. SCHOFIELD, WILLIAM. Research in clinical psychology: 1949. *J. clin. Psychol.*, 1950, 6, 234-237.
31. SCHOFIELD, WILLIAM. Changes in responses to the Minnesota Multiphasic Inventory following certain therapies. *Psychol. Monogr.*, 1950, No. 311, 64.
32. SEEMAN, WILLIAM. Clinical opinion on the role of therapist adjustment in psychotherapy. *J. consult. Psychol.*, 1950, 14, 49-52.
33. TUCKER, J. E. Rorschach human and other movement responses in relation to intelligence. *J. consult. Psychol.*, 1950, 14, 283-286.
34. WECHSLER, DAVID. *The Measurement of Adult Intelligence.* Williams and Wilkins, Baltimore, 3rd edition, 1944.
35. WEIDER, ARTHUR AND NOLLER, PAUL A. Objective studies of children's drawings of human figures: I. Sex awareness and socio-economic level. *J. clin. Psychol.*, 1950, 6, 319-325.
36. WITTENBORN, J. R. Level of mental health as a factor in the implications of Rorschach scores. *J. consult. Psychol.*, 1950, 14, 469-472.
37. WITTENBORN, J. R. Statistical tests of certain Rorschach assumptions: Internal consistency of scoring categories. *J. consult. Psychol.*, 1950, 14, 1-19.

THE PROGNOSTIC VALUE OF THE METENYM TEST IN A FOLLOW-UP STUDY OF PSYCHOSURGERY PATIENTS AND THEIR CONTROLS

JOSEPH ZUBIN AND CHARLES WINDLE

Psychiatric Institute, Columbia University

INTRODUCTION

The purpose of this study was to determine the value of the Metenym (multi-meaning word) Test in predicting outcome of psychosurgery. This test, previously used by Capps[2] under the name of Homograph Test, consists of a list of common stimulus words: (note, dock, cell, hatch, pole, round, quarter, file, club, band, diamond, gin) each of which has at least five different meanings. The subject is required to give as many meanings as he can for each word. The score is the total number of different meanings (excluding the first) and has been regarded as a measure of verbal flexibility. The data presented here have been previously reported in the Columbia-Greystone study[1, 5]. At that time follow-up information was not available, but by the end of the second year after operation it was possible to select sixteen patients with clearly defined outcome. Eight patients who stayed out of the hospital continuously since the time of discharge (out-group) were contrasted with eight who had remained continously in the hospital in an unimproved condition (in-group)[6]. Four patients in each of these two selected groups had been subjected to psychosurgery, the other four being controls.

DEVELOPMENT OF SCORING SYSTEM

A scoring system was developed similar to the one utilized in scoring the vocabulary of the Stanford Binet Test[11]. Acceptable responses included definitions listed

in the *Shorter Oxford English Dictionary*[8] and specific usages of these meanings. In the preoperative performance on the Metenym Test the average number of shifts in meaning (meanings in addition to the first) given by the in-group was 18.9, while the mean of the out-group was 14.0, a difference significant at the .10 level. A tabulation of the number of different meanings given for each stimulus word revealed that in ten of the words, the in-group exceeded the out-group, in one word the two groups were equal, while in the last word the out-group excelled. The probability of one group exceeding the other by chance alone in 10.5 out of 12 comparisons is less than .001, if the comparisons are assumed to be uncorrelated. It may be concluded that the in-group showed a higher capacity for shifting on this test than did the out-group. An item analysis of each word revealed that the critical point of differentiation between the two groups was the zero shift category (one response only), the incidence of non-shifting being more than twice as great in the out-group than in the in-group.

To achieve a minimum overlap between outcome groups, each word was weighted in accordance with the degree to which the frequency of zero shifts differentiated between the two criterion groups.* The weights given for the first shift were as follows: note—10; dock—8; cell, hatch and pole—4; and the remaining words—2. For each additional shift a weight of one half of that given for the first shift was credited. The total weighted shift score obtained was subjected to an analysis of variance in order to determine the relationship of the preoperative score to outcome and operative status. The prognostic value of the test was significant on the .05 level, but the difference between the operated and control groups was negligible.

Validity of the Scoring System

Before accepting the difference between the in- and out- groups as dependent upon basic ability on the Metenym Test, it is necessary to investigate incidental factors which may produce differences in performance on this test. A comparison of the frequency of similar or repeated response definitions indicated no differences in this respect between outcome groups. Differences in word knowledge or in tendencies to give responses of differing degrees of difficulty were investigated by means of the Semantic Word Count of Lorge and Thorndike[7]. No differences between outcome groups were found in tendency to give meanings of different levels of difficulty. The Metenym Test was found to correlate .67 with the Wechsler-Bellevue IQ. In order to determine whether the test derived any part of its prognostic strength from the IQ, the IQ level was partialled out by means of the analysis of covariance technique. The prognostic value of the Metenym Test remained on the .05 level.

Prognostication for Total Group

Our scoring system was developed empirically with criterion patients of clear-cut outcome. The next step was to apply the scoring system to all the patients in the Columbia-Greystone Project. This meant the addition of nineteen patients whose follow-up status was less well defined. They could, however, be classified into three groups: the in-group (unimproved), out-group (improved) aud out-in group (fluctuating status). The mean score of the total in-group was 49.8, the mean of the out-group was 32.6, while the out-in group mean was 35.7. In an analysis of variance it remained clear that the preoperative test score predicted outcome significantly regardless of operative status.

Postoperative Effects

An analysis of variance of the immediate postoperative test scores obtained within the first month after operation indicated no significant differences between outcome groups or between operated and control groups. However, when the preoperative differences between groups were eliminated by the analysis of covariance,

*This goal could have been obtained more directly by means of the discriminant function[4], but in view of the smallness of the sample, inspection was used to determine the proper weights.

a significant difference between the operated and control groups was revealed. Apparently the operation either reduces the ability in the Metenym test immediately after operation or interferes with the usual improvement expected from repetition of the test. The preoperative IQ level was found to be significantly related to this loss, the duller patients tending to show losses more often than the brighter patients ($P < .05$), especially within the operated group. It may be concluded that the tendency to lose in the immediate postoperative testing is most marked in the duller operated patients regardless of eventual outcome. The second and third postoperative tests were administered approximately three months and one year after operation respectively. For neither testing period did analyses of variance or covariance reveal significant differences between outcome or treatment groups.

Discussion

Although no cross validation of these findings is now available, the fact that the prognostic value of the test is not altered when the sample size is increased to include patients whose outcome was less clearly determined lends further credence to the reliability of the findings. There are, however, two aspects of the results which are somewhat unexpected and therefore demand careful scrutiny. First, why should the patients who do poorly on this test have better prospects of recovery than those who do well? Second, why should the prognosis hold equally well for the operated and the unoperated? One is tempted to provide theoretical speculation to account for these two findings, but until there is verification from cross validation studies, it is fruitless to theorize.

Summary

The Metenym Test was found in retrospect to be prognostic of outcome in a two year follow-up study of 19 operated and 16 unoperated cases. The prognostic score obtained empirically on selected criterion cases held up when applied to the entire group of patients, but showed no difference between treatment groups. Contrary to previous studies, good performance on this test was *not* indicative of *good*, but of *poor* prognosis. The immediate postoperative testing indicated a loss in the operated group, but the expected practice gain in the unoperated group. These postoperative changes bore no relationship to outcome. Later postoperative testings were unrelated to either outcome or treatment groups.

Bibliography

1. Beechley, R. M., and Rust, R. Other tests. Ch. 22 in [3].
2. Capps, H. Vocabulary changes in mental deterioration. *Arch. Psychol.*, 1939, *34*, No. 242.
3. Columbia-Greystone Associates, Mettler, F. A. (Ed.) *Selective partial ablation of the frontal cortex.* New York: Hoeber, 1949.
4. Fisher, R. A. *Statistical methods for research workers.* Edinburgh: Oliver and Boyd, 1938.
5. Landis, C. Test results one year after operation. Ch. 23 in [3].
6. Landis, C. in F. A. Mettler (Ed.) *The first conference on psychosurgery.* Nat'l. Inst. Mental Health, New York, Nov. 1949. Gov't Printing Office (In press.)
7. Lorge, I., and Thorndike, E. *A semantic count of English words.* New York: Columbia University, 1938.
8. Onion, C. T. (Ed.) *Shorter Oxford English dictionary.* Oxford: Oxford University Press, 1933.
9. Severn, D. Heteronyms for psychodiagnostic use. *J. consult. Psychol.* 1949, *13*, 144-145.
10. Sheer, D., and Shuttleworth, M. Psychometric studies. In Columbia-Greystone Associates, F. A. Mettler (Ed.) *Psychosurgical problems.* Phila.: Blakiston, 1951.
11. Terman, L. M., and Merrill, M. A. *Measuring intelligence.* Boston: Houghton Mifflin Co., 1937.

A STUDY OF THE RELATIONSHIP BETWEEN THE VERB-ADJECTIVE QUOTIENT AND THE RORSCHACH EXPERIENCE BALANCE

WILLIAM HAYS, SAUL GELLERMAN, WILLIAM SLOAN

Lincoln State School and Colony

Introduction

The present study purports to examine the possibility of a relationship between a sampling of an individual's speech and the nature of his Rorschach record. Specifically, we are here concerned with the possible relationship of the verb-adjective ratio to the ratio of M/C. Several workers have pointed out the relationship between personality "type" and speech usage. Some have inferred that Rorschach determinants and speech usage might be related as functions of the same dynamic process. Busemann[3], Stern[7], Boder[2], and Sanford[6] each expressed the view that the relative incidence of adjectives and verbs in individual speech bears some relationship to type of personality. Boder also suggested the possibility of similarity of speech type and Rorschach's extraversive and introversive "Experience-types". Balken and Massermann[1] found that the Verb-Adjective Quotient differentiated among three clinical types: conversion hysteria, anxiety neurosis, and obsessive compulsive neurosis. Furthermore, their results are strikingly analogous to the description of Experience-type for these three groups given by Rorschach[5], suggesting that relationship between speech usage and Rorschach movement and color responses might exist.

Although several promising leads have been suggested concerning a possible relationship between VAQ and Experience balance on the Rorschach, previous studies have been somewhat limited as to scope or in experimental design. The present study is an effort to investigate systematically the relationship of the VAQ to the M/C ratio over a wide range of types of subjects and to set up an experimental design which will minimize the operation of experimental errors. The usefulness of such a study is felt to lie in the possible utilization of the verb-adjective quotient as an additional instrument in the investigation of personality.

Procedure

The present experiment is designed to evaluate the following hypotheses:

A. The verb-adjective ratio in a controlled sample of individual speech will vary directly as the ratio of M to the weighted ΣC on the Rorschach record.

B. The verb-adjective ratio will vary directly as M/C.

C. The verb-adjective ratio will also vary directly as M, FM, M/C, Y, T.

D. These relationships will obtain in varying degrees in subject samples differing in levels of intelligence, as approximated by different IQ and/or CA relationships.

Several ratios of Rorschach determinants were chosen as representations of the general index of individual experience balance. The M/ΣC score (employing the weights arbitrarily assigned by Rorschach to various form levels of C) is the classical index of experience type. However, it was decided to employ an unweighted M/C ratio as well, in order to permit evaluation of any effects which the form-level weighting of C might produce. Finally, the ratio of unweighted $\frac{M+FM+m}{C+Y+T}$ was utilized, on the assumption that FM, m, Y, and T may be interpreted as but variations in the general processes represented by M and C respectively. Rorschach scoring followed the Beck method, with the exception of the FM and m determinants which were scored according to the Klopfer criteria.

A design based on analysis of covariance is suggested by these hypotheses. Accordingly, the experiment was designed to consist of a Rorschach examination and a

"speech sampling" in each of a number of groups differentiated according to mental and chronological age level. Thus, both the V/A and M/ΣC relationship and its variation with regard to level of intelligence could be evaluated.

Six experimental groups were selected to represent different gradations of IQ and/or CA. It should be emphasized at the outset that these groupings are not meant to constitute a random sampling of the general population, but rather a stratified random sample of specific IQ and CA levels, within the ranges appropriate for the Rorschach method. Table 1 summarizes the characteristics of the samples. Each group contains twelve subjects, (six males, six females), for a total sample of 72 subjects. As far as could be determined, no subject had had any previous experience with the Rorschach test.

TABLE 1. CHARACTERISTICS OF EXPERIMENTAL GROUPS

	Group	N	C. A. Range (in years)	I.Q. Range	Source
I	Adolescent Mental Defectives	12	below 17	50-65*	Patients, Lincoln State School and Colony
II	Average Children	12	9-10	102-111**	4th Grade Pupils, Public Schools of Lincoln, Illinois
III	Adult Mental Defectives	12	above 17	68-89***	Patients, Lincoln State School and Colony
IV	Adolescent Normals	12	15	100-106**	High School Sophomores, Public Schools of Lincoln
V	Average Adults	12	over 21	90-110***	Employee Applicants, Lincoln State School
VI	Superior Adults	12	over 21	over*** 120	Graduate Students (not Psychology Majors) Illinois State Normal University

*Stanford = Binet
**Otis
***Wechsler-Bellevue

The technique was as follows: Each subject was presented with a standard series of nine T.A.T. cards. These were, in order: 13B, 3BM, 12BG, 9BM, 15, 14, 19, 17BM, 18BM. These cards were selected for comparative simplicity of formal content. It was felt that no important advantage would derive from the inclusion of a larger number of cards. The usual T.A.T. instructions were employed and the responses of the subjects were wire recorded. The resulting transcribed protocol constituted the "speech sample". V/A ratios were derived by tallying the number of verbs and the number of adjectival forms in each speech sample. The criteria employed by Balken and Masserman[1] were adhered to in this tabulation. Also, in the present study forms of *to be* and *to have*, used either independently or as auxiliary verbs, were not counted.

The Rorschach test was administered in the usual manner. Two experienced examiners were employed, and, for each subject, both tests were administered by the same examiner in one sitting. The order of administration was alternated from subject to subject.

Results

Before these data could be subjected to statistical analysis, certain mathematical difficulties inherent in the manipulation of ratios had to be considered. Cronbach[4] points out that the distributions of ratios derived from Rorschach scores tend to be skewed, making statistical treatments assuming linearity invalid for these purposes. Furthermore, the incidence of zero scores in the numerator or denominator of the ratios renders their algebraic treatment difficult. Both objections may be obviated, however, through the logarithmic transformation of the data.* For this purpose the formula log (V+1/A+1) was employed for the V/A ratio, and log (M+1/ΣC+1) for the M/ΣC ratio. The analysis of covariance was undertaken with these converted logarithmic scores as data. Analysis of covariance was chosen for the treatment of these data as it affords a test of the significance of the differences in regression between several groups. Thus, while the F test for significance of differences among adjusted group means was made in each analysis (being insignificant in each case), the F test involving differences in group regressions is the one which is of interest in regard to the present hypotheses. Should this F be insignificant, it may be assumed that the relationship here in question does not vary to a significant degree between groups varying in IQ and/or CA relationship.

Therefore, the following statistical treatments were employed for each set of data: first, the over-all product-moment correlation between the two variables for all 72 subjects was calculated. Analysis of covariance then was used to determine the variances for subgroups. Finally, Bartlett's Chi-square Test was used to determine the homogeneity of variance for each variable between subgroups. Should both the analysis of covariance and Bartlett's test be insignificant, average correlation from the analysis of covariance could be accepted as the best estimate of population correlation.

As Cronbach points out, the validity of M/ΣC ratio as a personality indicator is in dispute. One cannot maintain, for example, that the implications of an M/ΣC ratio of 10/5 are the same as those of a ratio of 2/1. However, this study is concerned primarily with the mathematical relationship of this ratio with an external variable and not with individual psychological implications. Purely psychological objections to this procedure therefore seem unwarranted in this study.

Analysis of covariance and Bartlett's test were not significant for each relationship of ratios. Therefore, the average correlation from the analysis of covariance was accepted as the best estimate of relationship between each set of variables. The test involving log (V+1/A+1) and log (M+1ΣC+1) showed an experimental correlation of +.615 (significant beyond the 1% level), and an average correlation from covariance of +.644. The experimental correlation between log (V+1/A+1) and log (M+1/C+1) was +.591 (significant beyond the 1% level), with an average correlation of +.639. Finally, the experimental correlation of log (V+1/A+1) to log (M+FM+m+1/C+Y+T+1) was +.731 (significant beyond the 1% level), with an average correlation of +.744.

Discussion

The results of this experiment indicate, rather conclusively, a concomitance of variation between the verb-adjective ratio in verbalization and the M/C ratio on the Rorschach test. This correlation obtains without significant variation in all the various CA and IQ levels represented by the subjects in this study. These findings, by and large, are consistent with hypotheses advanced by earlier investigators, and suggest that both speech usage and Rorschach determinants are products of a characteristic of personality structure.

However, the possibility of a causal relationship between speech usage and Rorschach determinants operating in either direction should be pointed out. For

*The writers are indebted to Dr. Lee J. Cronbach of the University of Illinois for valuable suggestions regarding the statistical procedures employed in this study.

example, characteristically "adjectival" persons could be found to have an extratensive experience balance on the Rorschach because of the type of words predominantly employed.

On a more practical level, the present study points to the feasibility of controlled studies of verbal production and its relationship to personality characteristics. Thus, study of the larger area of speech content and of more Rorschach determinants may prove theoretically and clinically fruitful.

Summary and Conclusions

1. Seventy-two subjects, divided into six groups of twelve subjects each on the basis of chronological age and mental level, were administered the Rorschach and Thematic Apperception Test.

2. Verb-adjective ratios were computed from the T.A.T. stories of each subject, and were compared with each of three different estimates of the Rorschach experience-balance.

3. Analyses of covariance were carried out on logarithmic transformations of these data. Each of these analyses revealed a significant positive relationship between the variables. The subgroup correlations did not vary significantly in any analysis.

4. It is concluded that speech usage bears a definite relation to the nature of the Rorschach record. Some possible implications of this finding are discussed.

Bibliography

1. Balken, E. and Masserman, J. The language of phantasy: III. The language of the phantasies of patients with conversion hysteria, anxiety state, and obsessive-compulsive neuroses. *J. Psychol.*, 1940, 10, 75-86.
2. Boder, D. The adjective-verb quotient: A contribution to the psychology of language. *Psychol. Rec.*, 1940, 3, 309-344.
3. Busemann, A. *Die Sprache der Jungend als Ausdruck der Entwicklungsrhytmik.* Jena, 1925.
4. Cronbach, L. Statistical methods applied to Rorschach scores: a review. *Psychol. Bull.*, 1949, 46, 393-429.
5. Rorschach, Herman. The application of the form interpretation test, in *Psychodiagnostics.* (Ed. 2) Bern: Huber, 1932. (Translated by Lemkau, P., and Kronenberg, B.) New York: Grune and Stratton, 1942.
6. Sanford, Fillmore H. Speech and personality. *Psychol. Bull.*, 1942, 39, 811-845.
7. Stern, W. A Busemann: Die Sprache der Jugend. *ZPdPPs.* 1925, 26.

THE FIGURE-BACKGROUND RELATIONSHIP IN CHILDREN WITH CEREBRAL PALSY

JANE E. DOLPHIN, PH.D. AND WILLIAM M. CRUICKSHANK, PH.D.

Department of Psychology
Florida State University

School of Education
Syracuse University

Introduction

Studies have been undertaken and reported in which attempts were made to substantiate the findings of the studies by Werner and Strauss and also to determine whether or not characteristics of the perceptual disturbances which they noted among exogenous mentally retarded children would be observed in the performance of children with cerebral palsy [2, 3]. Throughout the research of Werner and Strauss, exogenous children (in contradistinction to endogenous mentally retarded children) were observed to have difficulty in determining the appropriate figure-background relationship in certain test situations to which they were exposed [8, 9]. Similarly, the present writers have observed that the performance of children who have cerebral palsy is significantly impaired on visual motor and tactual motor tasks due in part to the interference of the background structure. Because the accurate perception of the figure-background relationship is so important to the learning process, further examination of this factor was undertaken.

Method

Since the original purpose of the study had been to corroborate the general findings of other authors by using a population of cerebral palsied children, close adherence to earlier methodology was maintained [10]. In the present investigation two groups of thirty children were used as subjects, sixteen boys and fourteen girls in each group. The first group were children who represented all classifications of cerebral palsy and who had a mean chronological age of 10.03 years, a mean mental age of 9.5 years, and a mean regressed intelligence quotient [6] of 93.5. The second group consisted of physically normal children whose mean chronological age was 10.17 years, mean mental age, 9.5 years, and whose mean intelligence quotient was 93.7. To these two groups of children two tests involving the differentiation of the figure-background relationship were administered. Both tests were presented in somewhat modified form to that used by the earlier investigators.

The first test, *a picture test*, consisted of a series of nine cards on which were depicted black and white line drawings of objects, including a hat, a milk bottle and a cup, a knife, a boat, a chicken, an iron, a basket, and a hand holding a stick. These pictures were embedded, as shown in Fig. 1, in clearly structured backgrounds consisting of jagged and waving lines, squares, crosses, and other abstract forms. The pictures were individually tachistoscopically exposed for one-fifth of a second twice in succession. The child was requested to "Tell me what you see." The responses of the children were recorded verbatim.

The second test, *a multiple choice test*, like the picture test, was used to test the hypothesis that disturbances in figure-background relationship exist in the cerebral palsy children. In this test a geometrical figure, modified after Strauss and Werner [7] and constructed of heavy circular dots embedded in a configuration of small dots, was presented to the child for one-half second. Immediately after the presentation the child was shown a series of three cards. One of these cards contained only the background of the original card, the second showed the original background with a different geometrical figure embedded in it, and the third showed the original figure embedded in a different background. The child was asked to choose the card which seemed to him most like the test card as follows: "Pick out the one most like the one you just saw." The responses of the subjects were again recorded verbatim.

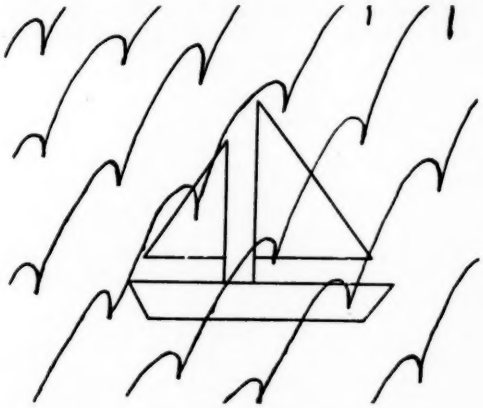


Fig. 1. Illustration of one of the nine cards presented tachistoscopically to the subjects.

Results

The Picture Test. The responses of the two groups of children to the picture test adapt themselves to five categories: (1) the correct figure response in which the child correctly names the figure with no mention of the background, (2) the incorrect figure response in which the child recognizes and names a figure which is similar in shape to the correct one, e. g., a bowl for a cup, (3) background with correct figure in which the subject correctly names the figure and also describes the background. (4) background with vague figure in which the subject correctly describes the background but only recognizes the figure as a vague shape, such as, a circle or a square, and (5) pure background responses in which the subject describes only the background with no mention of the figure.

Table 1 compares the responses of the cerebral palsy and normal groups in the ability to differentiate figure from background. It is seen that the cerebral palsy children are inferior to the physically normal group in differentiating the figure from the background in all but one area and that these differences are significant. In the ability to distinguish the figure from the background and give only the correct figure

Table 1. Comparison of frequency scores of cerebral palsy and normal groups for the picture test

Picture Test / Type of Response	Cerebral Palsy Group Frequency Scores	Normal Group Frequency Scores	Chi Squares	Level of Significance
Correct figure	16	74	44.86	.001
Incorrect figure	8	56	40.84	.001
Background with correct figure	55	53	.0378	.90
Background with incorrect figure	97	60	12.36	.001
Background	94	27	47.8126	.001

response, there is a better than .001 percent level of significance between the two groups. The cerebral palsy group has a frequency score for the *correct figure response* of 16; the normal group, 74.

There is a significant difference between the responses of the cerebral palsy and normal groups on the *incorrect figure responses.* The chi square between the two frequencies of 8 for the cerebral palsy group and 56 for the normal group is 40.84 giving a better than .001 percent level of significance.

The difference between the groups on the *background with correct figure response* is negligible. Here a 90 percent level of significance is derived from a chi square of .0378. The great similarity in the frequency of the two groups of subjects in this type of response might be expected because of the inability of the two groups to clearly define both fields and give one the preference over the other.

Other striking differences between the cerebral palsy and normal groups are in the *background with incorrect or vague figure* and *background* types of response. There is a significant difference at the .001 percent level between the responses of the cerebral palsy and normal groups on the *background with vague figure* type of response. The chi square between the two frequencies of 97 for the cerebral palsy group and 60 for the normal group is 12.36, giving a better than .001 percent level of significance between the groups. In the pure *background* type of response there is again evident a significant difference between the two groups with the chi square between the two frequencies of 94 for the cerebral palsy group and 27 for the normal group being 47.8126 and significant at better than the .001 percent level of significance.

The Multiple Choice Test. The responses that the two groups made to this test adapt themselves to three different categories: (1) the pure background response in which the subject selected the card with only the original background; (2) the original background with different geometrical figure response, and (3) the different background with original figure response.

Table 2 gives a comparison between the responses of the cerebral palsy group and those of the normal group in the ability to differentiate the figure from the background. It is seen that the cerebral palsy children are inferior to the physically normal children in differentiating figure from background although this difference is not significant except in one area. In the ability to select the original figure regardless of the background there is a better than 1 percent level of significance between the

TABLE 2. COMPARISON OF FREQUENCY SCORES OF CEREBRAL PALSY AND NORMAL GROUPS FOR THE MULTIPLE CHOICE TEST

Multiple Choice Test Type of Response	Cerebral Palsy Group Frequency Scores	Normal Group Frequency Scores	Chi Square	Level of Significance
Background	10	5	2.2219	.10
Original background with different figure	18	15	.6058	.50
Different background with original figure	2	10	6.6640	.01

two groups, the cerebral palsy group having a frequency response of 2; the normal group, 10. There is no significant difference between the responses of the cerebral palsy and normal group in the type of response involving the selection of the original background with different figure. The chi square between the frequency scores of 18 for the cerebral palsy group and 15 for the normal group is .6058 giving a better than 50 percent level of significance.

The difference between the responses of the groups in the selection of the background only with no figure recognition is not significant. The chi square between the frequency scores of 10 for the cerebral palsy group and 5 for the normal group is 2.2219 which is significant at the 10 percent level. However, the selection of the background only type of response more frequently by the cerebral palsy group is consistent with the results of the picture test which showed a figure background disturbance in the cerebral palsy children.

Summary

The statistical differences in the responses of the cerebral palsy and normal groups to both the nine situations involving figure-background relationships and to the items of the multiple choice test showed that the cerebral palsy children were inferior to normal children in distinguishing the figure from the background. Werner and Strauss [10] reported such disturbances as characteristic of the perception of the exogenous mentally retarded children. Halstead [4], Cotton [1], and Machover [5], in connection with research on other populations, have made somewhat similar observations. In part this difference might be due to the phenomena of forced responsiveness to extraneous stimuli which has been found characteristic of organic pathology. Further the presence of meticulosity also characteristic of individuals with cortical damage may account for the fact that when the cerebral palsy children were able to differentiate the figure they also felt compelled to include the background in their descriptions. The pathology of the figure-background relation in the cerebral palsy child has significant implications for educators and for educational methodology. Such will be discussed in detail in a later paper being prepared by the present writers.

References

1. Cotton, C. A. A study of the reactions of spastic children to certain test situations. *J. genet. Psychol.*, 1941, 52, 27.
2. Dolphin, J. E. and Cruickshank, W. M. Visuo-motor perception in children with cerebral palsy. In press.
3. Dolphin, J. E. and Cruickshank, W. M. Tactual motor perception in children with cerebral palsy. In press.
4. Halstead, W. C., and Settlage, P. H. Grouping behavior of normal persons and of persons with lesions of the brain. *Arch. Neurol. Psychiat.*, 1943, 49, 489.
5. Machover, K. *Personality projection in the drawing of the human figure.* Springfield, Ill.: C. C. Thomas, 1949.
6. Rolon, P. J. Problems of regression. *Harvard Educ. Rev.*, 1941, 11, 213.
7. Strauss, A. A. and Lehtinen, L. *Psychopathology and education of the brain injured child.* New York: Grune & Stratton, 1947.
8. Werner, H. Rorschach method applied to two groups of mental defectives. *Amer. J. Ment. Def.*, 1945, 49, 304.
9. Werner, H. and Strauss, A. A. Causal factors in low performance. *Amer. J. Ment. Def.*, 1940, 45, 213.
10. Werner H. and Strauss, A. A. Pathology of figure-background relation in the child. *J. abn. soc. Psychol.*, 1941, 36, 236.

A FACTOR ANALYSIS OF SOME SZONDI PICTURES[1]

WILSON H. GUERTIN

Ionia State Hospital[2]

INTRODUCTION

The Szondi test employs six pictures in each of eight different diagnostic categories. All six of the pictures within a given category are regarded as having a common significance. No distinction is made among the pictures in a given category and the only test observation made is the total number of the pictures within a category that are selected. This procedure implies that there are strong loadings of different group-factors for each category.

One would expect the pictures in a given category to be highly correlated with one another, but only slightly correlated with pictures from other diagnostic categories. And, of course, a factor analysis should reveal eight group factors, each closely associated with a different diagnostic category. Since Szondi test construction is based upon the assumption that an important group-factor underlies each of the diagnostic categories, this hypothesis will be investigated. A centroid factor analysis should disclose eight group-factors, each corresponding to a different diagnostic category.

DESIGN

Subjects. A total of 100 subjects with personality and/or behavior difficulties was used. An equal number of each sex was employed to provide as wide a range of personality variables as possible. In addition, a broad age range from 15 to 71 years was used. All subjects were predominantly from the lower socio-economic class. Three different populations were drawn upon to comprise the whole experimental sample of 100.

(1) Twelve male delinquents from a correctional institution. None of these was clearly psychotic or neurotic; hence, they were regarded as behavior problems. The most frequent delinquency was some form of stealing.

(2) Twenty female mental hospital cases. Diagnoses varied but paranoid schizophrenia seemed to predominate.

(3) Sixty-eight were patients at a hospital for the criminally insane. The diagnoses were varied but schizophrenia was predominant. The crimes committed by these patients also varied in nature and seriousness. There was a rather heavy loading of murderers and sex offenders in this group.

Pictures. The two Szondi pictures from each diagnostic category which had been shown to have the highest preference values (2) were used.[3]

Procedure. The subjects were asked to rank all the pictures in their order of preference from high to low. These results were recorded on McBee type sorting cards for each subject in terms of "liked" (rank 1-8) or "Disliked" (rank 9-16). The necessary frequencies were obtained from these cards and tetrachoric correlation coefficients

[1]The author wishes to express his appreciation to the many who have made this investigation possible: Dr. Albert I. Rabin for his encouragement and suggestions relating to experimental design; Drs. Leo Katz and James Karslake for statistical procedure; and to the Department of Public Welfare of the State of Michigan and the superintendents of the Boy's Vocational School at Lansing, Kalamazoo State Hospital, and Ionia State Hospital for providing the subjects. Finally, special acknowledgment is made of the support and encouragement of Dr. P. C. Robertson, Superintendent of the Ionia State Hospital, where this project was executed.

[2]Now at Beatty Memorial Hospital, Westville, Indiana.

[3]Sixteen other pictures of "normal" Americans, described elsewhere (2), were included among the pictures presented to the subject but are not to be considered in this paper.

were estimated from Thurstone's charts[1]. The assumption that these rankings are distributed normally has to be made, but no particular objection to this was apparent.

One of the Szondi pictures (Vh) had to be excluded from the final matrix because the frequency distribution along the continuum was too uneven to yield reliable coefficients. The poorest distribution among the remaining pictures comes from a cutting line at 29 per cent of the cases for one picture.

The factor analysis of the Szondi intercorrelations was obtained by the complete centroid method of Thurstone[3, p. 161 ff]. Communalities were estimated by selecting the highest correlation in each column. Two trial factorings before the final factoring made it possible to adjust these estimates of communality to the final values.

Results

Table 1 gives the factor loadings for the individual Szondi pictures studied. The last column gives the communalities extracted (sum of the squared factor loadings). Four group-factors seem to emerge from the analysis quite clearly. A fifth one and perhaps others exist but the large error in these coefficients prevents the extraction of more than five. Almost half the variance of the intercorrelation matrix can be attributed to these five extracted factors. Much of the remainder is undoubtedly error variance.

Table 1. Complete Centroid Factor Matrix of the Szondi Pictures (decimal points omitted)

Szondi Picture	Factors I	II	III	IV	V	Sum a_{ij}^2
VIh	27	30	–43	–15	–37	51
IIs	–31	–24	–30	–33	26	42
IVs	–43	–29	20	10	–33	43
IVe	15	19	47	30	14	39
Ie	51	47	–38	–12	06	64
IVhy	79	16	–17	26	17	78
Vhy	44	18	33	–23	02	39
IVk	25	13	47	17	06	33
VIk	–64	50	–38	26	–12	89
IVp	–18	–30	38	47	15	51
Vp	–14	34	21	–22	06	23
IId	28	–41	–21	30	06	38
Vd	51	–25	–11	45	10	55
Vm	–47	28	–27	31	06	47
IIm	–28	38	18	45	14	48
Sum a_{ij}^2	2.6	1.5	1.5	1.3	0.4	7.40

The nature of the data discouraged more refined treatment such as rotation of axes and identification of factors. The factor matrix presented is that obtained by the complete centroid method and each factor is orthogonal or independent of each other.

It would be expected according to the design of the Szondi test that the factor loadings would be more similar for pictures of the same diagnostic categories than for pictures of different categories. The mean difference in the squared factor loadings, retaining the original sign, for the pictures of similar category is 13.7 (SE 9.88). The mean difference for the factor loadings of pictures in dissimilar categories is 14.5. The difference in obtained values is slightly in the anticipated direction but is wholly attributable to chance. The final residual matrix shows that most of the common factors have been extracted. There are a few high correlations left; in order of magnitude they are .47, .31, .30 and .29. The mean absolute coefficient in this residual matrix is .085 with an SD of .08.

Discussion

In comparing the factorial constitution of the Szondi pictures within the same diagnostic categories, it is seen that they are as different as are pictures from different categories. It would appear that there is as much difference in meaning between pictures within the same diagnostic category as there is between pictures from different categories. This finding is quite at variance with Szondi's assumption that pictures within a diagnostic category have a similar meaning or diagnostic significance. There is no support for the present practice of regarding pictures within a similar diagnostic category as equivalents and merely observing the total number of selections within each diagnostic category.

The formal factor analysis discloses five group-factors which explain the obtained correlation matrix reasonably well, considering the error inherent in the dichotomous correlations. The factor matrix suggests that several of these factors frequently are important in determining the preference for a Szondi picture. Heavy loadings of a single group-factor are absent for all but one picture and the communalities of the individual pictures require the use of several factors to explain them. There is no support for Szondi's view that a single factor can explain the preference for a picture adequately.

It is encouraging to learn that the correlation matrix cannot be described adequately by one or two factors. It would appear that many factors are disclosed by the preferences for Szondi pictures. These preferences are probably determined primarily by personality variables since it is difficult to imagine other individual differences exerting appreciable effects. A test for evaluating personality should disclose many underlying factors since personality is viewed as multi-dimensional. Had the complete Szondi set been studied, other factors very likely would have appeared. The factor matrix suggests that it would be difficult to establish pictures that have heavy loadings of a single factor. Such "pure" test items are desirable but not essential in evaluating the individual differences in personality. It would seem that the most profitable approach in evaluating personality by picture preferences would be a multiple correlation one.

No attempt will be made to identify and describe the factors obtained because such a procedure is too uncertain with such a small number of variables with minimal loadings. Such might be possible if a factor analysis of all 48 pictures were made. In conclusion, it will be noticed that once again the Szondi test results have failed to conform to deductions made from the theory which the test construction subsumes. However, the general method of evaluating personality by picture preferences seems to show some promise.

Summary and Conclusions

1. Szondi test construction is based upon the assumption that an important group-factor underlies each of the diagnostic categories. An intercorrelation study of the pictures should disclose eight group-factors, each corresponding to a different diagnostic category. Since Szondi test results depend upon the stimuli, it is assumed implicitly that these pictures offer the optimum number of stimulus cues.

2. One hundred psychiatric patients of mixed diagnoses ranked two pictures from each of Szondi's diagnostic categories in the order of their preference. Tetrachoric intercorrelations between the pictures were calculated.

3. The complete-centroid factoring produced five group-factors sufficient to account for almost 50 per cent of the variance of the intercorrelation matrix.

4. The factor constitution of pictures of the same diagnostic category are no more similar than that of pictures from different categories. Hence, there is no support for the present practice of regarding pictures within the same diagnostic category as equivalents and merely observing the total number of such selections made within each category.

5. It would seem that the most profitable approach in evaluating personality by picture preferences would be a multiple correlation technique since pictures depending upon only one factor (pure items) are not readily seen.

6. The Szondi test results failed to conform to deductions made from the theory determining the test construction. However, the general method of evaluating personality by picture preferences seems to show promise.

References

1. Cheshire, L., Saffir, M., and Thurstone, L. L. *Computing diagrams for the tetrachoric correlation coefficient.* Chicago: Univ. of Chicago Bookstore, 1933.
2. Guertin, W. H. A comparison of the stimulus value of Szondi's pictures with those of normal Americans. *J. clin. Psychol.*, 1951, 7, 163-166.
3. Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univ. of Chicago Press, 1947.

SEX DIFFERENCES IN THE ROSENZWEIG P-F STUDY, CHILDREN'S FORM

GEORGE SPACHE

Reading Laboratory and Clinic
University of Florida

Introduction

The Rosenzweig Picture Association Method for Assessing Reactions to Frustration (P-F Study) is a projective device for evaluating the individual's modes of response to everyday situations of conflict. It permits the individual to identify with and respond for anonymous figures in a series of 24 cartoon-like drawings.[1] We have used it extensively in the study of elementary-school children and found that with certain modifications in the scoring that we have suggested[2], it is highly serviceable in revealing how the child handles his aggressive impulses and the modes of response he may take toward adults and other children.

Our experiences have led us to the belief that there are probably significant differences between boys and girls in their modes of handling aggressive impulses as revealed by this test. Certainly the roles of male and female conduct our children are induced to assume would, in our opinion, tend to produce differences in the expression of these impulses. For example, society expects much more obviously aggressive behavior in boys than in girls and, at the same time, is disturbed by an inversion of this behavior. This study was undertaken as a pilot test of the hypothesis of sex differences among elementary-school children as revealed in their responses to the P-F Study.

Method and Results

The test was administered to 50 boys and 50 girls. In most cases, the procedure was on an individual basis with the child dictating his responses to the examiner. We have found this method superior to group administration particularly among children under the fourth grade level since the younger pupils are severely handicapped in writing their answers because of their lack of writing vocabulary or fluency, and spelling difficulties. The characteristics of the separate sex groups and the entire population are compared with those of Rosenzweig's normative group in Table 1.

TABLE 1. COMPARISON OF EXPERIMENTAL GROUPS WITH ROSENZWEIG'S NORMATIVE GROUP IN TERMS OF PERCENTAGE OF RESPONSES

	Boys		Girls		Total Group		Norm Group		C. R. of differences	
N	50		50		100		256			
Age	6–13		4–13		4–13		4–13		Boys vs. Girls	Total vs. Normal
Factors	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.		
E	52.9	11.6	49.7	17.5	50.9	15.5	46.0	15.6	1.10	2.71
I	22.1	6.4	22.2	7.7	22.2	7.1	25.6	9.7	.07	3.83
M	26.5	9.5	29.4	13.0	27.9	11.4	28.5	11.0	1.31	.42
O-D	13.5	6.3	15.1	5.1	14.3	5.8	16.3	6.5	1.45	2.85
E–D	51.3	11.1	54.8	13.9	52.6	12.6	56.4	10.6	1.00	2.75
N-P	36.9	11.3	33.0	12.8	34.9	12.3	27.2	10.2	1.69	5.53
GCR	61.3	10.8	61.6	12.0	61.5	11.5	62.1	10.6	.13	.50

When compared with Rosenzweig's normative group, the present group is more aggressive (*E*) and less insightful or inclined to recognize its own faults (*I*); and tends to direct its aggressiveness toward the solution of conflict situations to a markedly greater degree (*N-P*), rather than venting it on the situation (*O-D*) or other individuals (*E-D*). The differences in *E* and *I* are probably due to the lesser age of our group, inasmuch as the trends shown here in these two categories are characteristic of younger children. Our group has a mean age of 9.3 with 36 per cent of the children ten years of age or older compared with the mean age of 9.9 for Rosenzweig's group where 54 per cent were ten years of age or older. This influence of the age factor is apparent in a comparison of our means with those given for various age groups.[1] Our group means closely resemble the norms for the 6-7 year old group in *E* and *I*, while Rosenzweig's group mean, influenced by the greater proportion of older children, resembles the 8-9 year old group in *E* and the 10-11 year old group in *I*. The differences in *O-D*, *E-D* and *N-P* are not explicable on the basis of age. They may readily indicate that the type of behavior characteristic of high *N-P* individuals, those persistent in trying to find solutions to their conflicts with others, tends to cause teachers to refer these children to the school psychologist for individual study.

The fact that our data include a number of children referred for problems of academic difficulty and personality adjustment may account for the greater aggressiveness and persistence as well as lesser maturity and self-control indicated in the group Rosenzweig results. Despite the fact that our experimental group does not resemble the population on which norms for the P-F Study were established, we may still make a comparison of the responses of boys and girls and note similarities and differences. The group means of Table 1 fail to show any truly reliable differences between the sexes in the proportions of responses of various types or categories. Comparisons of modal answers to specific items or attitudes toward the parent figure or child figure in the conflict situation may also be made however.

We used Rosenzweig's method of finding the significance of the difference between the modal answers. This technique selects a modal answer when the C. R. of the difference between one type of answer and the next most frequent is three or more. If differences are less than this significant level, two or three modal answers, depending upon the distribution of choices, may be designated. Using this method, we found the modal answers for our groups.

The similarities among our groups and Rosenzweig's normative population are much more marked than the differences. Of Rosenzweig's 33 modal answers, our total group agrees with 25, adds 13 others and disagrees with or drops eight. There is a tendency for our group to show a greater diversity of answers within the group

and therefore more modal answers than was true of Rosenzweig's population. In terms of particular items there are identical modal answers for eleven and marked disagreement only in items 10 and 23. Item 10 depicts a mother expressing regret for having had to send the child to bed. Many of our respondents were uncertain whether the child was being punished or whether it was merely bedtime. This uncertainty or misinterpretation resulted in a much greater degree of impunitiveness or acceptance of the situation (M^1 and M) among our group and less intropunitiveness or acknowledgment of probable fault (I) than was true in the normative group. Apparently we should have emphasized the element of punishment when children expressed doubt about the incident. Item 23 depicts a mother expressing regret that the boy's soup is cold. Our group showed less acceptance of the situation as such (M^1 and M) and a greater tendency to respond personally in admitting fault (I) or to reproach the mother (E).

The differences between our modal answers and the normal are the expected ones in view of the group tendencies shown earlier. The greater aggressiveness is shown in the tendency to add modal answers of the E and e type, boys exhibiting ten such modal answers not found in the norms; girls showing eleven such tendencies. In the N-P category, boys add nine new modal answers and girls add six.

Only one really significant difference between sexes is present in item 18 in which boys show a reliably greater response of the e type than do girls. This item depicts a situation in which a girl informs a boy that she is not going to ask him to her birthday party. Apparently this situation evokes persistence or attempts to insist upon the invitation more frequently among boys. It should be noted that the common response for our children was a combination of M^1 and E, a "sour grapes" or belittling response such as "I didn't want to come anyway". This type of response was found in 45 per cent of our population and is significantly more common than the single M^1 answers.

As we have pointed out in an earlier study [2], responses to the P-F Study may be separated into those reflecting adult-child and child-child attitudes. Thus the 13 situations depicting conflict between a child and an adult may be scored separately from the 11 situations involving child-child conflict. The responses of our sex groups may be separated similarly to determine whether there are sex differences in the expression of aggressiveness toward adults or toward children. These data are presented in Table 2.

TABLE 2. SEX DIFFERENCES IN ADULT-CHILD AND CHILD-CHILD CONFLICTS*

	Boys vs. Girls C.R. in A-C	Boys vs. Girls C.R. in C-C	C.R. Boys A-C vs. C-C	C.R. Girls A-C vs. C-C
E	+1.50	+1.10 ✱	3.58 (C-C) ✱ ✱	2.08 (C-C) ✱ ✱
I	.62	+ .21	13.7	12.15
M	+ .29	+1.3	1.89	.85
O-D	+ .21	+1.4	3.81	1.40
E-D	4.93	6.27	3.5 (C-C)	1.24 (C-C)
N-P	1.3	.28	9.7	7.38
GCR	.76	+1.2	2.22	.13

✱ + indicates girls exceeded boys in number of responses.

✱ ✱ Responses were greater in A-C than in C-C except in E and E-D, for both boys and girls.

*The means and standard deviations in terms of the number of responses to the P-F on which these critical ratios are based will be supplied by the author upon request.

Our data indicate that there are strong tendencies for girls to show greater degree of obvious aggressiveness (E) toward both adults and children. Our girls also show a greater degree of permissiveness (M) as well as greater projection of aggress-

iveness upon the environment (*O-D*) when in conflict with other children. Boys show a reliable tendency to react to the personalities involved, to direct their responses toward the other person (*E-D*) when in conflict with either adults or children. Our boys also show a stronger tendency to try to find solutions for conflict with adults. Finally our girls show a greater tendency to conform to the normative modal answers particularly in their conflicts with other children.

Intra-group comparisons of the attitudes toward the adult figure and the child figure are also possible. Thus we may compare the strength of the tendencies to give each type of responses when in conflict with an adult or with another child, as in Table 2.

Both boys and girls appear to differ significantly in their attitudes toward adults or other children. Both show tendencies to exhibit more obvious aggressiveness (*E*) toward other children than toward adults; more readiness to acknowledge blame (*I*) when in conflict with adults; and more tendency to show persistence in seeking solutions (*N-P*) when dealing with adults. In addition to these common trends, boys also show strong tendencies to project their aggressiveness against the environment when dealing with adults but against the other individual (*E-D*) when in conflict with children. Boys tend to show greater conformity when dealing with other children than with adults.

Certain other facts may also be elicited from our data. We noted the frequency of "defensive" or rationalizing answers commonly indicated by $\underline{E}$ or $\underline{I}$. These are evaluated as suger-ego responses. Some of these were sufficiently frequent as to assume the proportions of modal answers. For example, using the same criterion for the selection of a modal answer as previously, $\underline{E}$ in item 12 and 19, and $\underline{I}$ in item 22 are modal answers in our group. There are no reliable differences between sexes in these answers.

Conclusions

Comparisons of modal answers typical of boys and girls and the proportions of answers in the various categories have failed to show sex differences in our group of 100 elementary school pupils. These group comparisons based on the scoring of the P-F Study as a whole did not support the hypothesis of sex differences in the type or direction of aggressive impulses.

When conflict situations involving adult-child relationships are scored separately from those of child-child conflicts, sex differences do appear. We find boys showing a marked tendency to react personally and to direct their aggressiveness at the other individual when in conflict with adults or children. Girls show significantly less of this sensitivity or defensiveness. Boys show significantly more outward aggressiveness toward other children than toward adults; both sexes are more ready to acknowledge or accept blame and try harder to find peaceful solutions when in conflict with adults than with children. Boys also tend to project their aggressiveness upon the environment when in conflict with adults but towards the other individual when dealing with children.

Expressed in dynamic terms, these findings indicate that we may expect to find more frequent and more personal conflict between boys and their classmates than is true for girls. Children of both sexes defer to adult demands to a marked degree by admitting blame, apologizing or trying to find peaceful solutions that are acceptable to the adult. Boys are more conscious of the need for curbing obvious aggressiveness and avoiding personal conflict when dealing with adults. They seem to be aware of the fact that adults are quick to detect and punish obvious aggressiveness toward adults when committed by boys.

Bibliography

1. Rosenzweig, Saul. *The Children's Form of the Picture-Frustration Study.* 4562 Scott Avenue, St. Louis 1, Mo.: The Author.
2. Spache, George, Differential scoring of the Rosenzweig P-F Study, Children's Form. *J. Clin. Psychol.*, 1950, 6, 406-408.

SYMBOLIC MEANINGS IN THE RORSCHACH CARDS: A STATISTICAL STUDY[1]

EPHRAIM ROSEN

University of Minnesota

INTRODUCTION

In recent years there has been a trend toward the increasing use of content analysis of the Rorschach as opposed to purely statistical analysis. Two hypotheses that have been widely employed in content analysis are that: (1) responses to card IV can be interpreted on the assumption that the card is perceived by the subject, for the most part unconsciously, as a father image; (2) card VII can similarly be considered as symbolizing a mother figure. Associated with these hypotheses have been others: card X has sometimes been designated the "family card" and card VI has been called the "male sexuality" card. Investigation by the writer has revealed few explicit statements in clinical literature of these hypotheses. Although relevant discussions can be found in Bochner and Halpern [1] and Lindner [3], application of the hypotheses has spread, in the main, by informal communication between clinicians. Thus Dr. Bruno Klopfer has stated, in a personal communication to the writer,

" . . . Card IV is to a majority of subjects a symbol of masculine aggression which may or may not be related to the father image, and Card VII a symbol of basic emotional security which is frequently associated with the mother image. The fact that children between 4 and 6 years of age frequently see a 'house' in this area and that this is repeated by adults who have an unusually strong desire for basic security seems to reinforce the latter hypothesis. The association of a phallus with the top of Card VI is too frequent a phenomenon to be called a mere hypothesis." The last of the above hypotheses—that of a relation between card VI and male sex symbols—has been explicitly formulated and attacked experimentally, for example in a recent study by Pascal, Ruesch, Devine, and Suttell[4].

The present study attempted to test these hypotheses by a questionnaire which asked subjects to choose, from among the ten Rorschach cards, that card which most nearly brought to mind the feeling or association of a male sex organ, masculine aggression, authority, father symbol, mother symbol, and family symbol, respectively. Direct questioning calls into play processes not identical with those involved in unconscious, non-reportable, perception and symbolization. But it was felt that as a first empirical approach to the problem a naive questionnaire method might well be employed. Further, it was hoped that forced choice of a card for a given item would yield results not completely independent of symbolic and unconscious mental processes, yet more readily analysable by statistical procedure than a set of standard Rorschach protocols.

PROCEDURE

Three groups of subjects, all students in psychology classes and all without previous exposure to the Rorschach, were employed in the study. Groups 1 and 2 viewed a slide, in color, of the ten Rorschach cards projected simultaneously. Group 3 individuals were questioned one at a time. For this group the cards were presented simultaneously on a table, in the standard position. For all three groups the number of each card was affixed in the lower right hand corner.

All groups were given an answer sheet which contained the following instructions:

"The ink blots you see are the ten blots of the Rorschach test. It has been found that people see different things in the different blots, and that the blots remind

[1]The present research was completed before the publication of Meer and Singer's report on a similar problem. See Meer B. and Singer, J. "A note on the 'father' and 'mother' cards in the Rorschach ink blots." *J. consult. Psychol.*, 1950, *14*, 482-484.

people of various things and arouse different feelings. Whatever an ink-blot reminds you of is the correct response for you. There are no right or wrong answers.

"The following questions are asked in order to increase our knowledge of where people are most likely to perceive various things.

"The questions ask you which blot looks *most* like some specific thing to you, or which blot most nearly brings to mind some specific feeling or association. Some of the questions may seem far-fetched to you. Nevertheless, please choose *one*, and only *one*, blot as your answer to *each* question. You may perceive the suggested response in a whole blot, or part of a blot; it is up to you whether to use the whole or part. In either case, simply write the number of the blot. The same blot may or may not be your answer to a considerable number of different questions.

"In answering the questions you will probably do best to work rapidly, using your impression after scanning all ten blots for the suggested thing, feeling, or association."

On the basis of experience with Group 1, the experimenter added the verbal statement for Groups 2 and 3 that a choice need not be made if the subject felt that he absolutely could not make one. Consequently, there was a slightly larger number of unanswered questions in these two groups than in Group 1.

The questionnaire proper, on the reverse side of the sheet from the instructions, read "Which *one* blot most nearly suggests to you . . . "; there followed a list of 20 items after each of which a card number was to be written. The items directly relevant to the hypotheses of this study were scattered among the 20 in order to disguise the purpose of the questionnaire, and also to yield some information on the validity of the questionnaire method employed by noting whether expected cards were chosen for known popular[2] and near-popular responses. The 20 items were, in order: a bat, a house, human beings, animals, a male sex organ, a female sex organ, nighttime, human internal organs, a feeling of masculine aggression, an X-ray picture, a feeling of emotional insecurity, an ornamental design, a feeling of emotional security, a feeling of authority, clouds, a father symbol, a mother symbol, a family symbol, a landscape seen from a distance, and an animal skin.

Questionnaires were anonymous, except that Group 1 subjects who expected to be members of a class to be taught by the writer in the following academic quarter were asked to put some identifying symbol on their papers. Since Group 2 subjects were drawn from the second class a number of subjects, 7 male and 5 female, were members of both Groups 1 and 2. The total number of subjects in the study is presented in Table 1.

TABLE 1. DISTRIBUTION OF SUBJECTS IN THE THREE GROUPS

Group	Male	Female	Total
1	59	46	105
2	28	15	43
3	31	14	45

RESULTS

On a chance basis, each Rorschach card would be chosen by ten per cent of the subjects for each item of the questionnaire. Therefore, significance of association between card and item was evaluated by computing significance of difference between obtained percentage and the theoretical chance percentage of ten. Since the figure for chance percentage was not lowered to correct for occasional omission of responses by some subjects, any error is in the conservative direction.

Table 2 summarizes card choices by males, females and both sexes combined, for the three groups of the study, at confidence levels of 1 per cent, .1 per cent, and .01 per cent.

TABLE 2. CARD CHOICES: MALES, FEMALES, AND BOTH SEXES COMBINED

Item	Card	Group	Confidence Level		
			Male	Female	Both Sexes
1. Bat	V	All	.0001	.0001	.0001
2. House	VIII	3	.01	—	.01
	X	1	.0001	.0001	.0001
		2	.01	.0001	.0001
3. Human beings	II	1	—	.0001	.0001
		2	—	.01	—
	III	All	.0001	.0001	.0001
	VII	1	—	—	.001
4. Animals	II	1	.0001	.01	.0001
	VIII	3	.0001	—	.0001
	IX	1	.0001	.0001	.0001
		2	.0001	.01	.0001
5. Male sex organ	IV	1	.0001	.0001	.0001
	VI	1	—	—	.01
		2, 3	.0001	—	.0001
6. Female sex organ	VI	All	—	.0001	—
		1, 2	.0001	—	.0001
	IX	3	.0001	—	.01
7. Nighttime	I	1	.0001	.0001	.0001
		3	—	—	.01
	IV	1	.01	.0001	.0001
		2	.0001	.01	.0001
		3	.0001	—	.0001
	V	3	—	.01	—
8. Human internal organs	VIII	1, 2	.0001	.0001	.0001
	IX	1	.0001	—	.0001
		2	.0001	.0001	.001
		3	.0001	—	.0001
	X	1	—	.001	.001
		3	.001	—	.01
9. Masculine aggression	I	1	.01	—	—
		3	.01	—	.01
	IV	1	.0001	.0001	.0001
		2	.0001	.001	.0001
10. X-ray	I	2	.01	—	.001
		3	—	.01	—
	IV	3	.0001	—	.0001
	VII	1	.01	—	—
	VIII	1	.01	.01	.001
11. Emotional insecurity	VII	1	.01	—	—
	X	1	.0001	.0001	.0001
		2	.0001	—	.0001
		3	.0001	.001	.0001
12. Ornamental design	III	1	—	.001	.001
	VIII	2	.01	—	.01
	X	1	.01	.0001	.0001
		2	.01	.01	.0001
		3	.0001	—	.0001
13. Emotional security	II	1, 2	.0001	—	.0001
	VII	2	—	.01	—

TABLE 2—*Continued.* CARD CHOICES: MALES, FEMALES, AND BOTH SEXES COMBINED

Item	Card	Group	Confidence Level		
			Male	Female	Both Sexes
14. Feeling of authority	I	1	.001	—	.001
		3	.01	—	.01
	IV	1	.0001	.0001	.0001
		2	.0001	.001	.0001
	VI	3	.001	—	.01
15. Clouds	VII	All	.0001	.0001	.0001
16. Father symbol	IV	1	.0001	.0001	.0001
		2, 3	—	.01	.0001
	VI	3	.01	—	
17. Mother symbol	VII	1	—	.01	.01
		2	.01	—	.001
18. Family symbol	I	3	.01	—	.01
	X	1	—	.0001	.0001
19. Landscape	VII	1	.01	—	—
	IX	3	.01		.01
	X	1	.0001	.0001	.0001
		2	.0001	.001	.0001
20. Animal skin	IV	1	.0001	.0001	.0001
		2	.01	.0001	.0001
		3	.001	—	.01
	VI	1	.0001	.0001	.0001
		2	.0001	.001	.0001
		3	.0001	.0001	.0001

Analysis of the table reveals that on the whole the "mother," "father," and "male sexuality" card hypotheses have been borne out. Card IV was associated with "masculine aggression," and "feeling of authority" by two of the three groups, both sexes combined, at a .01 per cent level of confidence. It was associated with "father symbol" by all three groups at a .01 per cent level. Card VII was associated with "mother symbol" by two of the three groups. Breaking these results down by sex, it can be seen that the relationships still hold, although with weaker agreement between groups and at poorer confidence levels. Card VI was preferred for "male sex organ" by all groups, both sexes combined; and by male subjects in two of the three groups. Male and female subjects of Group 1 chose card IV, in which the major stimulus projection is in a downward direction. In the other two groups no significant choices by females emerged for this item. The "family card" interpretation of card X received some rather weak support: only Group 1 females and both sexes combined chose card X.

The hypotheses can, however, be viewed in another light, by noting *all* the cards which were chosen, at a significant confidence level, for a given item, or by noting all the items associated with a given card. For example, not only card IV, but also cards I and VI, were preferred for the "authority" item by a number of groups. Similarly, card IV was chosen at a significant level by one or more groups, both sexes combined, for "male sex organ, nighttime, X-ray," and "animal skin," in addition to its choice for "father symbol," "masculine aggression" and "feeling of authority." Card VI was chosen by one or more groups of both sexes combined for "female sex organ," "feeling of authority," and "animal skin," in addition to the

"male sex organ" choice. It may thus be said that while the hypotheses under test in this study have been supported, evidence has also emerged that no one card approaches identity in meaning for a great number of subjects. Different subjects can perceive a father figure in various cards, and the same card has various connotations for different subjects. Consideration of the actual per cent of subjects who made a given choice points to the same conclusion: for both sexes combined in Group 1, the largest group of the study, the only choice approaching unanimity was that of card V for "bat." The figure for this choice was 90.5 per cent. Card VII was designated for "clouds" by 66.7 per cent, and card VI for "animal skin" by 63.8 per cent. All other choices fell below 50 per cent. For example, card IV was chosen for "feeling of authority" by 35.2 per cent of this group. The relationships hypothesized in this study were thus borne out, but not with the strength evidenced for such popular or quasi-popular responses as "bat," "clouds," and "animal skin." In passing, it should be noted that the high percentage of subjects choosing expected cards for the popular and near-popular responses can be viewed as a partial check on the validity of the method used in the study.

Certain other interesting relationships emerge from the table. Card X was often chosen for "emotional insecurity." This finding can be interpreted as coherent with the hypothesis that many subjects experience color shock in response to colored cards. It is, of course, equally plausible that the choice was determined by confusion arising from the great numberof details in the cards. "Emotional security," on the other hand, was associated with cards II and VII. It is not clear what stimulus aspects of these cards motivated the choices. The choice of VII for "emotional security" by Group 2 females does not seem to have been associated with perception of the card as a "mother symbol," since females of this group showed no card preference for "mother symbol."

Extent of agreement between the three groups of the study can be analyzed for cross-validation purposes. Analysis of Table 2, for example, shows that of 23 choices at a .01 per cent level for both sexes combined of Group 1, 18 of the 23 choices were repeated in Group 2, and 11 in Group 3, at 1 per cent or better. Similar results emerge from analysis of data for the sexes separately: in general, agreement was far above a chance level, and greater between Groups 1 and 2 than between Group 2 and either of the other two groups. However, the strongest choices found in Group 3 *were* replicated in Groups 1 and 2: of the 12 Group 3 choices at .01 per cent, both sexes combined, 10 were repeated in Group 1, and 10 in Group 2, at 1 per cent or better. The inclusion of 12 of the same subjectsin both Groups 1 and 2 would thus seem not to have spuriously raised agreement between these groups to any great degree. On the other hand, the effect of differences in distance from the stimuli in the three groups does seem to have effected subjects' choices to specific questionnaire items. Group 1, numerically larger than Group 2 and therefore at the greatest average distance from the slide, tended to prefer cards II and IX for "animals"; Group 2 preferred card IX only; and Group 3, closest to the stimuli, preferred card VIII. Thus decreasing distance correlated with choice of smaller details. However, results for the questionnaire responses central to the hypotheses of the study seem not to have been affected in any direct fashion by distance from the cards and consequent difficulty of perception of details.

A partial reliability check was performed by comparing first and second responses of the 12 subjects to whom the questionnaire was administered twice. On a chance basis two choices would be repeated for the twenty questionnaire items. Analysis showed an average of 8.25 agreements between the two administrations. For the "masculine aggression," "feeling of authority," and "father symbol" choices, agreement was higher: the average number of agreements was 1.67, the maximum possible being 3.0, and chance 0.3. For the "mother symbol" item, there was agreement in 8 of the 12 choices. These agreements are all significant at the 1 per cent level, so that reliability appears to have been substantially better than chance.

SUMMARY

A questionnaire procedure was used to test the hypotheses that cards IV, VI, VII, and X are perceived by subjects as representing a father symbol, male sexuality, a mother symbol, and a family symbol, respectively. Subjects were asked to designate the card most nearly bringing to mind the feeling or association of these items and a number of others. Results indicated considerable support for the first three of these hypotheses, in terms of per cent of three separate groups choosing the predicted cards. The "family symbol" hypothesis alone received but little support from the data. However, the results also indicated great individual differences in perception of symbolic meanings of Rorschach cards, so that no single meaning was assigned to a specific card by a majority of any group. The Rorschach would thus appear to consist of stimuli which have a partial, but not a total, symbolic communality for subjects.

REFERENCES

1. BOCHNER, R. and HALPERN, F. *The clinical application of the Rorschach Test.* New York: Grune & Stratton, 1942.
2. KLOPFER, B. and KELLEY, D. M. *The Rorschach technique.* New York: World Book Co., 1946.
3. LINDNER, R. Content analysis in Rorschach work. *Rorsch. Res. Exch.*, 1946, *10*, 121-129.
4. PASCAL, G. R., RUESCH, H. A., DEVINE, C. A. & SUTTELL, B. J. A study of genital symbols on the Rorschach test: presentation of a method and results. *J. abn. soc. Psych.*, 1950, *45*, 286-295.

THE RELATIONSHIP BETWEEN CERTAIN MEASURABLE FUNCTIONS OF AUTONOMIC NERVOUS SYSTEM ACTIVITY AND COLOR RESPONSES ON THE RORSCHACH TEST*

HALCYON HUGHES, LEON J. EPSTEIN, AND HUDSON JOST

University of Tennessee

There is clinical evidence to support the hypothesis that color responses on the Rorschach test are related to affective tone [12, 14, 15, 24, 26, 28, 29,] but little experimental evidence. The present analysis represents an attempt to study this relationship quantitatively through the use of continuous measurements of respiration, relative blood pressure, heart rate, and galvanic skin resistance as measured by the Keeler polygraph during the free association period of the Rorschach test. There appears to be ample experimental evidence [1, 2, 9, 11, 13, 17, 21, 22, 23, 25, 31, 32, 33, 34] which would seem to indicate that there is a relationship between the physiological measures under investigation and so-called emotional or affective states.

Attempts have been made to correlate physiological measurements with various psychological tests of personality and attitude [5, 6] but few have involved their correlation with scoring categories on the Rorschach test, and fewer still have used the method of obtaining these measurements simultaneously with the administration of the test.

Brower [4] in 1947 correlated certain Rorschach categories from the protocols of subjects selected at random from a college undergraduate population, with measurements of diastolic blood pressure, pulse pressure, and heart rate which were taken

*From the Psychophysiological Laboratory of Gailor Psychiatric Hospital, Department of Psychiatry, College of Medicine, University of Tennessee, Memphis. The authors are indebted to Dr. T. S. Hill, Director of the Hospital, and to Mr. William J. von Lackum, Chief Clinical Psychologist, for their interest and their helpful suggestions. This study was made possible through a grant-in-aid from the Life Insurance Research Fund.

before and after a state of visuo-motor conflict. Such a conflict was induced by means of a mirror-drawing test, under the conditions of direct vision, mirror vision, and the blindfolded state. He found that higher pulse rate prior to the experiment and higher pulse pressure after the experiment tend to accompany personality traits of constrictiveness and repressiveness, and concluded that cardiac excitation has some concomitant relation with those personality trends indicated by "F" in the Rorschach test. Furthermore, it appeared that subjects showing lower levels of adjustment to reality, as measured by the "FC" Rorschach category, in Brower's work, exhibited higher diastolic blood pressure prior to the experiment. He concluded also that the greater the "residual extraversial reserve" (responses to VIII, IX, X divided by the total number of responses), the greater the tendency toward high pulse pressure and low diastolic blood pressure, after taking the test as a whole.

Frost and Rodnick [10] examined normal and schizophrenic subjects in an effort to discover the relationship between certain Rorschach determinants and concomitant galvanic skin responses. Normal subjects showed significantly larger deflections with form-dominant responses than did the schizophrenic subjects. Form, in Rorschach terminology, indicates control [3], and the authors concluded that the schizophrenics seem to have an altered capacity to control and modify their reaction to reality.

Rockwell, Welch, Kubis, and Fisichelli [27] studied the relationship between galvanic skin resistance and "color shock" on the Rorschach test. Continuous graphic recordings were made of variations in palmar skin resistance during the administration of the Rorschach test to both normal and clinically psychoneurotic subjects. The results were treated statistically on the basis of each subject's galvanic skin response score per card. This score resulted from the sum of the skin resistance changes for each response on the card, as calculated by the ratio between the basal resistance at the beginning of each response, and the actual change in resistance as measured by the height of each curve. The clinical group presented fewer responses as well as a lower mean level of skin conductance changes.

A recent study involving measurement of the galvanic skin responses accompanying the Rorschach test is reported by Levy [19, 20]. Fifty male college students were examined individually and their maximum galvanic deflection per card was measured. There was no significant increase in skin conductance accompanying the colored plates and Levy, therefore, concluded that Rorschach cards do not differ among themselves in affective value. This conclusion was based upon the assumption that increased skin conductance indicates affect. Oddly enough, however, Levy found for her sample a marked change in skin conductance associated with the position of the card in the series. When the Rorschach cards were systematically presented in a rotated order, the eighth in the series, no matter which card it happened to be, was accompanied by galvanic changes that were significantly beyond the one per cent confidence level. Levy's experiment differs from the present study, however, in several important aspects. For the sake of uniformity, Levy modified the routine instructions which precede a Rorschach examination. Her subjects were directed to give only one response to each card. Furthermore, the maximum galvanic deflection *per card* was measured, whether or not the response concerned included the element of color ("C") in its scoring. The purpose of Levy's study was to determine the relative affective value of each Rorschach card and her study, therefore, is not comparable to the present analysis in which the aim is to compare the relative value of individual responses.

Method

The subjects of this study were thirty-two students from the Division of Medical Sciences of the University of Tennessee, selected by choosing every second person from an alphabetical list. Fifteen of the subjects were student nurses and seventeen were medical students, the total group ranging in age from twenty to thirty-five

years. The Rorschach test was administered individually, by qualified examiners, in accordance with the technique suggested by Beck [3]. The testing situation was entirely new to every subject. The examinations were conducted without interruption in a quiet and private office with none present other than the subject and two examiners, one to operate the polygraph not in view of the patient and one to administer the Rorschach test. Since diurnal variations have been observed in the galvanic skin response [18], all examinations were given in the morning. The light source in every case was daylight plus fluorescent lighting [3, p. 3].

The blood pressure cuff was placed on the subject's right leg, just below the knee, and the galvanic electrodes were placed on the left foot. These devices were placed on the legs rather than on the arms, as is the customary procedure, in order to leave the hands entirely free for normal handling of the Rorschach cards. An earlier unpublished study by Jost [16] indicates that measurements taken in this manner from the lower extremities give results comparable to those obtained from the arm and hand. At regular intervals, after Rorschach cards III and VI, the polygraph was stopped and pressure reduced in the cuff in order that good circulation might be restored in the subject's leg.

At the beginning of each examination, after adjusting the polygraph, a control run of two minutes duration was made to determine the initial heart rate, respiration, relative blood pressure, and skin resistance of the subject in the resting state. During the administration of the test, the experimenter operating the polygraph pressed the stimulus marker whenever the subject made a response, and made a written note on the moving record in order that the polygraph recording could be matched with the Rorschach protocol. A final control run was taken at the end of the Rorschach free association period. No recordings were made on the polygraph during the inquiry.

The entire series of Rorschach cards was administered in the standard order. However, since the purpose of this study was to demonstrate the relationship between color responses and so-called affective physiological changes, only responses to the Rorschach cards containing color, namely II, III, VIII, IX, and X, entered into the calculations. Within these cards, physiological changes obtained simultaneously with those responses containing an element of color in their scoring were compared with physiological changes obtained with those responses containing no color element.

Results

The results of this study are presented in Table 1. Physiological measurements were eliminated, for the purpose of this study, on those Rorschach responses made so close together that it was impossible to determine to which response the measurements were associated. Measurements of respiratory movements during speech were also not included. Because of these omissions, the number (No.) of measurements reported under the various categories in Table 1 is variable.

A comparison of the average respiratory rate recorded before color responses with that before non-color responses reveals very little difference in rate. The same is seen to hold true for a comparison between the average respiratory rate after color and that after non-color responses. These results also reveal little discernible difference in average rate of respiration between color and non-color Rorschach responses, measured either fifteen seconds before or fifteen seconds after the response. A comparison of the average amplitude of respiration before color and before non-color responses reveals no appreciable difference between these means. Following the responses, however, there seems to be a slight tendency for respiration to be less deep in the case of non-color responses. The average deviation of respiratory amplitude was used as a measure of respiratory stability. It is interesting to note that this deviation became smaller following the non-color responses, and larger following the color responses.

TABLE 1. PHYSIOLOGICAL CHANGES OCCURRING WITH COLOR AND NON-COLOR RESPONSES TO THE RORSCHACH CARDS (N = 32 CASES).

		COLOR			NON-COLOR			t
Measure		No.	M	S. D.	No.	M	S. D.	Color: Non-Color
RESPIRATION								
Frequency	Before	52	4.2	1.0	102	4.2	1.1	0.24
	After	51	3.8	1.1	98	3.9	1.9	0.18
Amplitude	Before	51	8.7	5.3	102	8.8	5.3	0.07
	After	50	9.8	6.1	97	8.1	4.5	1.76
Stability (Av. dev. of amp.)	Before	52	2.4	1.5	103	2.9	2.4	1.59
	After	50	2.8	2.3	96	2.4	1.8	0.92
GALVANIC SKIN RESISTANCE								
Absolute change, in 1000 ohms		121	1.0	1.1	291	1.1	1.2	1.03
Per cent of total resistance		121	1.4	1.6	291	1.6	1.9	0.09
Per cent recovery in 15 seconds		120	31.5	40.5	291	38.5	42.7	1.38
HEART RATE								
15 seconds before response		171	90.5	17.1	405	89.7	16.5	0.52
During response		165	92.5	17.1	394	91.8	17.8	0.29
15 seconds after		165	90.5	17.0	398	89.9	16.7	0.38
BLOOD PRESSURE								
Maximum change								
	Systolic	125	3.6	2.7	316	3.7	2.2	0.15
	Diastolic	122	4.3	2.9	304	3.9	2.4	1.10
Time in sec. before R. that BP began to change								
	Systolic	125	1.6	2.4	315	1.2	1.9	1.74
	Diastolic	122	2.1	2.7	303	1.6	2.1	1.93
Duration in sec. of BP change								
	Systolic	111	7.0	3.6	282	7.0	4.6	0.18
	Diastolic	113	7.4	3.6	273	7.6	4.6	0.41
Per cent recovery in 15 seconds								
	Systolic	125	82.3	33.0	314	77.3	37.8	1.19
	Diastolic	122	86.4	27.6	301	78.3	36.6	2.07

The changes in galvanic skin resistance presented in Table 1 indicate a tendency for galvanic deflections to be slightly greater on the non-color responses than on the color responses. "Recovery" from the change was measured by the extent to which the galvanic skin response deflection returned, in fifteen seconds, to its pre-response level. In the amount of recovery, the tendency was toward a greater recovery on the non-color responses than on the color responses. These results tend to confirm Darrow's statement that the larger the galvanic reaction, the quicker the recovery [7]. Thus, these results, although differences are not significant statistically, indicate a trend toward a larger increase in galvanic deflection plus a greater recovery for the non-color responses.

The general picture of heart rate is one of marked consistency. A slight increase in average rate of not more than two beats per minute over the pre-response rate was noted during verbalization. In the case of both color and non-color responses the rate returned to the pre-response level within fifteen seconds after the response.

Table 1 presents the data regarding maximum blood pressure changes within fifteen seconds before each response. Color responses appeared to initiate greater mean diastolic pressure changes than did the non-color responses, but the changes are of rather low magnitude both in the case of systolic and diastolic pressure measurement. The differences in time, before color and non-color responses, that the systolic and diastolic pressures began to change, is presented in Table 1. These changes seemed to occur somewhat sooner in the case of color responses and furthermore the measure appears more variable in the case of color responses. Table 1 also reveals that, on the basis of the duration of the blood pressure changes, color and non-color responses could not be differentiated with respect to either systolic or diastolic blood pressures. With respect to recovery from the change in blood pressure, there is a tendency for recovery to be more marked in the case of color responses, particularly with respect to diastolic pressure.

Summary and Conclusions

The purpose of this study was to investigate the quantitative relationship between color responses on the Rorschach test (a supposed indicator of affect) and certain measurable functions of the autonomic nervous system which have been associated with affective tone. The principal conclusion reached on the basis of this study was that color and non-color Rorschach responses were not associated with appreciably different physiological changes as measured in this study. In other words, the so-called "affective" and "non-affective" Rorschach responses were not differentiated physiologically by the techniques employed in this study.

Bibliography

1. Abel, Theodora M. Attitudes and the galvanic skin reflex. *J. exp. Psychol.*, 1930, *13*, 47-60.
2. Armstrong, Harry G. The blood pressure and pulse rate as an index of emotional stability. *Amer. J. med. Sci.*, 1938, *195*, 211-220.
3. Beck, Samuel J. *Rorschach's test. I. Basic processes.* New York: Grune and Stratton, 1944.
4. Brower, Daniel. The relation between certain Rorschach factors and cardiovascular activity before and after visuo-motor conflict. *J. gen. Psychol.*, 1947, *37*, 93-95.
5. Brower, Daniel. The relations of visuo-motor conflict to personality traits and cardiovascular activity. *J. gen. Psychol.*, 1948, *38*, 69-99.
6. Cofer, Charles N. The psychogalvanic response as an indicator of emotional reaction to personality test items. *Amer. Psychologist*, 1948, *3*, 303. Abstract.
7. Darrow, C. W. The significance of the galvanic recovery curve. *Psychol. Bull.*, 1931, *28*, 678. Abstract.
8. Darrow, C. W. The relation of the galvanic skin reflex recovery curve to reactivity, resistance level, and perspiration. *J. gen. Psychol.*, 1932, *7*, 261-273.
9. Duffy, Elizabeth, and Lacey, Oliver L. Adaptation in energy mobilization: changes in general level of palmar skin conductance. *J. exp. Psychol.*, 1946, *36*, 437-452.
10. Frost, C. F., and Rodnick, E. H. The relationship between particular Rorschach determinants and the concomitant galvanic skin responses for schizophrenic and normal subjects. *Amer. Psychologist*, 1948, *3*, 277. Abstract.
11. Gaskill, H. V., and Cox, G. M. Patterns in emotional reactions: II. Heart rate and blood pressure. *J. gen. Psychol.*, 1941, *24*, 409-421.
12. Goldstein, Kurt. Some experimental observations concerning the influence of colors on the function of the organism. *Occup. Ther. and Rehab.*, 1942, *21*, 147-151. Cited by Rickers-Ovsiankina (26).
13. Greenwald, D. Ulrich. Individual differences in electrodermal response to continuous affective stimulation. *Univ. Ia. Stud. Psychol.*, No. XX, 1936, *48*, 632-646. Cited by Levy (19).
14. Hertz, M. R., and Baker, E. Personality patterns in adolescence as portrayed by the Rorschach ink-blot method: II. The color factors. *J. gen. Psychol.*, 1943, *28*, 3-61.
15. Janesch, E. Gefuhl und empfindung (Emotion and perception). Ber. 15. Kongr. dtsch. Ges. Psychol., Jena, 1936, 65-70. *Psychol. Abstr.*, 1938, *12*, 84. Abstract No. 768.
16. Jost, Hudson. Personal communication. December 1948.
17. Landis, Carney. Psychology and the psychogalvanic reflex. *Psychol. Rev.*, 1930, *37*, 381-398.
18. Landis, Carney. Electrical phenomena of the skin. *Psychol. Bull.*, 1932, *29*, 693-751.

19. LEVY, JEANNE R. Changes in the galvanic skin response accompanying the Rorschach test. *Amer. Psychologist*, 1948, *3*, 335. Abstract.
20. LEVY, JEANNE R. Personal communication. Changes in the galvanic skin response accompanying the Rorschach test. Paper read before the Fortieth Annual Meeting of the Southern Society for Philosophy and Psychology, March 1948.
21. MARSTON, W. M. Sex characteristics of systolic blood pressure behavior. *J. exp. Psychol.*, 1923, *6*, 387-419.
22. MORRIS, DON P. The effects of emotional excitement on pulse, blood pressure, and blood sugar of normal human beings. *Yale J. Biol. Med.*, 1935, *7*, 401-420.
23. MORRIS, DON P. Blood pressure and pulse changes in normal individuals under emotional stress; their relationship to emotional instability. *Psychosom. Med.*, 1941, *3*, 389-398.
24. OESER, O. Some experiments on the abstraction of form and color. Part I. Tachistoscopic experiments. *Brit. J. Psychol.*, 1932, *22*, 200-215. Cited by Rickers-Ovsiankina (26).
25. RACKLEY, L. E. The blood pressure and galvanic reflex as indicators of emotional states. *J appl. Psychol.*, 1930, *14*, 497-504.
26. RICKERS-OVSIANKINA, MARIA. Some theoretical considerations regarding the Rorschach method. *Rorschach Res. Exch.*, 1943, *7*, 41-53.
27. ROCKWELL, F. V., WELCH, L., KUBIS, J., and FISICHELLI, V. Changes in palmar skin resistance during the Rorschach test. *Monthly Rev. Psychiat. Neurol.*, 1947, *113*, 129-152.
28. RORSCHACH, HERMANN. *Psychodiagnostics.* Trans. by Paul Lemkau and Bernard Kronenberg. Berne, Switzerland: Hans Huber, 1942. New York: Grune and Stratton.
29. RUESCH, J., and Fensinger, J. The relation of the Rorschach color response to the use of color in drawings. *Psychosom. Med.*, 1941, *3*, 370-388.
30. SCHACHTEL, E. G. On color and affect: contributions to an understanding of Rorschach's test. II. *Psychiatry*, 1943, *6*, 393-409.
31. SCOTT, JOHN C. Systolic blood pressure fluctuations with sex, anger, and fear. *J. comp. Psychol.*, 1930, *10*, 97-114.
32. SHOCK, N. W., and COOMBS, C. H. Changes in skin resistance and affective tone. *Amer. J. Psychol.*, 1937, *49*, 611-620.
33. WECHSLER, D., CRABBS, L. M., and FREEMAN, R. Galvanic responses of pre-school children. *J. genet. Psychol.*, 1930, *38*, 203-222.
34. WHITEHORN, J. C., KAUFMAN, M. R., and THOMAS, J. M. Heart rate in relation to emotional disturbances. *Arch. Neurol. Psychiat.*, 1935, *33*, 712-731.

A BEHAVIOR RATING SCALE SUITABLE FOR USE IN MENTAL HOSPITALS[1]

RUBEL J. LUCERO AND BILL T. MEYER

Fergus Falls State Hospital, Minnesota

INTRODUCTION

The staffs in mental hospitals have long needed an objective, quantifiable measure of the behavior of the many patients who are inaccessible to existing psychometric and psychiatric methods. These people, mute, unintelligible, hyperactive or seclusive, make up approximately fifty percent of a hospital population, and are material for the various types of therapies including lobotomy, insulin shock, electro-convulsive, etc. The psychologist in his research role has attempted to evaluate the effectiveness of these therapies, but he has been blocked when present psychometric methods were of no avail. The system of rating a patient by qualitative statements ranging from *worse* to *complete recovery* is relatively valueless for scientific study.

This problem has not been studied too intensively in the past. A few attempts have been made, however, and they will be briefly discussed here. Malamud et al.[3, 4, 5, 6] devised a scale concerned largely with dynamics, and got a high degree of agreement among three psychiatrists rating twenty patients (r's of .868 and .907). Several factors make this scale impractical for general use in mental hospitals. First the areas to be rated (judgment, sexual, autism, etc.) are such that a maximum amount of training is necessary for raters. Then, the amount of time necessary to make these observations precludes its use on any large scale studies. Although it is a five point scale, it does not lend itself to quantification because, in actuality, it is a three point scale measuring in two directions, i.e. activity is rated from underactivity to normal activity to overactivity. Cohen et al[1] at Norwich State Hospital made up a five point scale for use by nurses and attendants on disturbed wards. The experimental validation was carried out with two nurses who had previous experience in rating patients, and correlations of .65 to .90 were obtained in six areas of activity. The main criticism is that too small an area of behavior is rated; however, the scale appears to be adequate within its limitations.

Other studies are being carried on; however, no data have been published on them, therefore they will be discussed only briefly. Janis[2] mentions a questionnaire used to evaluate Electro-Shock Therapy, but it appears that it is rather qualitative and that the patient has to be in contact sufficiently to answer the questions. Wittman[7] stresses the need for a quantifiable measure and has devised a scale measuring 26 areas of behavior. Criticisms of the scale are that the scoring method is complicated, and the language used and the areas rated (Delusional Ideation, Paranoid Projections, Insight, etc.) are obviously only for highly trained and experienced raters.

Analysis of the problem resulted in the formulation of the following five requisites for an adequate behavior rating scale:

1. It should be usable by relatively untrained raters, psychiatric aides (attendants). Any training necessary to achieve optimal reliability should be minimal.
2. The ratings should be such that they can be made in a very short time.
3. The scale should have quantifiable results.

[1]We are indebted to William Schofield for the original idea and for his many helpful criticisms; to Reuben Silver for his help with the experimental design and statistical critique; to John Pearson for his advice and for helping to make the project possible; to William L. Patterson for authorizing the project; and to all the psychiatric aides without whose cooperation the project would have been impossible.

4. It should measure an adequate sample of behavior that can be objectively rated.

5. It should be sensitive enough to pick up differences in behavior between different patients and the changes in behavior in one patient over a period of time.

With these objectives in mind, the L-M Fergus Falls Behavior Rating Scale was developed. The basic form of this scale was a rating scale devised by William Schofield for use on the psychiatric ward of the University of Minnesota Hospitals. The original form was revised seven times to fulfill the above requirements. Certain areas were eliminated, others added, and throughout an effort was made to clarify and simplify the definitions. An analysis was made of each revision in order to find the area of greatest disagreement, which was then revised. Through this process the final form evolved and is herein reproduced along with a copy of the instruction sheet. (Table 1 and Table 2).

TABLE 1. L-M FERGUS FALLS BEHAVIOR RATING SHEET

Name Hospital Number Age Sex Marital Status Date

Diagnosis Time on Present Ward Ward Religious Preference

A. Work
- ___Does no work—refuses—extremely negativistic.
- ___Does a little work with a lot of urging. Constant supervision is necessary.
- ___May have a regularly assigned job—and supervision may be necessary.
- ___Enthusiastic participation in all types of work—asks for work.
- ___Normal interest in work—ie., interested in some kinds of work more than others (will do other kinds than main interest if called upon to do so.)

If works, at what, and how many hours a day.....

B. Response to meals
- ___Has to have special attention, as eats too much, spoon fed or tube fed.
- ___Eats by self, is sloppy—may need coaxing.
- ___Eats by self using knife, fork and spoon properly. May show some finickyness.
- ___Passes and asks for things to be passed, but will not carry on table conversation.
- ___Would not stand out among normal people for eating habits.

C. Response to other patients.
- ___Stays all alone or may strike out at other patients.
- ___Will be with other patients only for a short while and with urging.
- ___Some signs of friendliness—speaks to patients—may have a friend.
- ___Some spontaneity in making contacts with other patients. May initiate play or work of a social relatively high order type. (Card game, washing dishes).
- ___Helpfullness expressed toward other patients—or non-hostile recognition of their being mentally ill and making allowances.

D. Response to Psychiatric Aides and nurses.
- ___Negativistic—hostile (can include striking)—doesn't do anything requested.
- ___Will do a few things if asked or pushed—shows no open hostility.
- ___Will do most things when asked—will ask for simple things—"I want my toothbrush".
- ___Extremely cooperative—will do anything when asked.
- ___Normal give and take relationship. Speaks spontaneously to nurses about things of not immediate importance. (weather, baseball games, etc.)

E. Response to Doctors, Social Workers, Psychologists.
- ___Hostile
- ___Passively negativistic (would rather not have anything to do with them but will not resist).
- ___Will speak when spoken to.
- ___Seeks advice.
- ___Understand, accepts, and asks for therapy.

F. Response to Electric or Insulin Therapy.
- ___Hostile—etc.
- ___Anxious, apprehensive, but not overly hostile.
- ___Passively accepts.
- ___Accepts positively—(May say, "I feel better after").
- ___Asks for, understands necessity for.

TABLE 1. L-M FERGUS FALLS BEHAVIOR RATING SHEET—(Continued)

G. Occupational therapy and recreational therapy (Walks don't count.)

___Does not participate at all—negativistic—hostile.
___Participates with urging for short periods.
___Participates when asked—some spontaneity.
___Shows interest—participates in all types wholeheartedly without discriminating very much between different types—looks forward to.
___Interested in many varied activities—normal selectivity (Likes some kinds more than others).

What patient is most interested in.

.....

H. Attention to dress and person.

___Has to be dressed—needs special attention of one kind or another.
___Dresses self but is sloppy.
___Some interest in looks—too much lipstick, fairly neat.
___Cares about looks and dress. Will ask for makeup or shaving equipment inconsistently. (Not an overall balance).
___Normal (for culture)—would not stand out in a crowd.

I. Psychomotor Activity (Not including going to the bathroom or meals).

___Stays in one place unless pushed, or hyperactive (seclusion necessary, etc.)
___Moves around a little (one chair to another) or if hyperactive, the activity is not of a type making seclusion or other restrictions necessary.
___Some activity resulting from the influence of the illness (moves around because voices say to) and some purposeful behavior.
___Behavior mainly purposeful. Still moves around a little fast or a little slow.
___Normal activity—would not stand out among normal people.

(If hyperactive or stays in one place, which one?)

.....

J. Speech

___Mute or speaks a lot but it doesn't make sense.
___A few words that make sense ("yes" or "no").
___Speaks in short, clear sentences, "Can I have my toothbrush".
___Speaks normally except a little fast or slow.
___Speaks normally.

(If mute or senseless talk, which one)

.....

K. Toilet Behavior

___Untidy anytime during the day and/or more than twice a week nightly.
___Untidy once or twice a week nightly—brushes teeth and washes only when told to do so.
___Not untidy—toilet behavior somewhat sloppy—brushes teeth and washes once a day without being told.
___Toilet behavior normal except for being too neat or too much time spent at one thing or occasionally sloppy.
___Toilet behavior normal.

L. List below anything physically wrong with patient's arms or legs.

.....

.....

M. List below anything physically wrong with patient's hearing, sight and speech organs.

.....

.....

N. List below any physical illness that patient has.....

.....

.....

TABLE 2. INSTRUCTIONS FOR FILLING OUT BEHAVIOR RATING SHEET

At the top of the Behavior Rating Sheet fill in as much information as can be found on the ward. Below you will find eleven different groups of descriptions of a particular type of behavior. Place a check (X) by the description that comes closest to telling how this particular patient has behaved for ..... If the patient's behavior has changed in this period, then check the description which tells how the patient has behaved *most* of the time. If the patient's behavior is evenly divided, then check two descriptions (or more, if necessary). Look at each group separately, *do not try to give an overall impression at any time* because, as you well know, a certain patient may be a very good worker but he may not speak at all. If you feel that you don't know enough about a certain bit of behavior (example, you haven't seen the patient eat) then *leave that part out*, it is far more important to have true descriptions than to have many descriptions. There will be other descriptions that you will have to leave out. (Examples: if patient isn't getting insulin or electric treatments, you can't rate that particular thing. If patient is bed ridden he (or she) can't be rated for amount of activity.) There are blanks under some of the groups. Fill them in as per instructions.

Procedure

Two inexperienced raters (psychiatric aides) from each of seventeen wards of Fergus Falls State Hospital rated independently three patients selected randomly from the raters' wards. Eight other inexperienced raters rated twenty-eight regressed and lobotomized patients who were part of a *total push* program. All raters were given the instruction sheet shown above as the sole instruction. There were eleven areas which could be rated, and numerical values were assigned to the five points in each area. These values ranged from one for the poorest type of behavior to five for the best type of behavior. Any difference over one point was considered to be a significant difference as is customary on a five-point scale. By inspection of Table 1, it can be seen that differences of three or more points are too divergent to be regarded as actual differences of opinion about a patient's behavior in a particular area. Thus ratings with differences of three or more points were considered to be invalid.

Results of reliability study. The raters took less than five minutes to rate each patient. Thirty-four raters making 866 ratings on fifty-one patients, representing every type of ward in the hospital, agreed 90 percent of the time. When invalid ratings were excluded (out of 866 ratings, 38 were invalid), the agreement was 95 percent. A breakdown of these figures showed that male raters agreed 89 percent, excluding invalid ratings (out of 394 ratings, 12 were invalid) 92 percent; female raters 92 percent, and excluding invalid ratings (out of 472 ratings, 26 were invalid) 97 percent. The eight raters making 462 ratings on 28 patients on a specialized program agreed 95 percent. There were no differences between male and female raters, and the number of invalid ratings was so small that the percentages were unaffected. It may be noted here that the majority of the invalid ratings were made by two raters who have difficulty in reading and writing. Although it is felt that percent agreement is a more satisfactory statement of relationship between raters than correlation, others have stated their results in terms of correlations, so, in order to compare results, correlations were computed. Ratings by four raters (two male, two female) on 20 patients of a *total push* program were subjected to small sample statistical treatment. Because of the small N, rho was computed in place of Pearsonian r, and small sample statistics were used throughout. The results were as follows:

Male raters: rho equals .92; M, equals 2.6, M_2 equals 2.6; S.D., equals 1.02, $S.D._2$ equals .95; S.E. $M._1$–M_2 equals .154; t equals 0.00; F equals 1.14.

Female raters: rho equals .94; M_1 equals 1.8, M_2 equals 1.9; $S.D._1$ equals .33, $S.D._2$ equals .54; $S.E._{M_1-M_2}$ equals .077; t equals 1.299 (1.383 required for significance at 20 percent level); F. equals 2.65. An F of 3.23 is required for significance at the five percent level; therefore the differences are not significant, and it may be assumed that the raters are in essential agreement. The scale apparently has fairly high reliability and objectivity.

Results of validity study. The scale has face validity (e.g. speech, dress are directly observable). With high reliability and objectivity, the significant difference in ratings of a patient's behavior over a length of time can be assumed to be due to a change in behavior, thus indicating the instrument is valid. This is borne out by a study on 10 patients of a *total push* program. The mean rating at the beginning of the program was 1.6, and six months later nine patients had ratings about the *pre-total push* mean, and one patient remained at the mean. At the beginning of the program all ten patients had ratings below the *post-total push* mean, which was 2.3. The *t* for the difference was 4.861 which is significant beyond the .1 percent level.

Other evidence of validity was obtained from an analysis of the ratings obtained from the total hospital sample. Averaging the ratings of wards of each type, a very interesting difference was noted. Although the sample from each ward was too small to make the evidence conclusive, it is nevertheless directly in line with assumed differences in wards. Regressed wards M equals 1.9; disturbed wards M equals 2.5; average wards M equals 3.0; geriatric wards (well-behaved seniles) M equals 3.5; convalescent and worker wards M equals 3.7; and the admission wards M equals 4.0.

The scale appears to fulfill reasonably well the requirements previously mentioned, and should prove a valuable addition to present psychometric tools.

Limitations and technical suggestions. Despite the high reliability, it is felt that an average rating from two raters is preferable to a single rating. The scale could lend itself as well to projects concerned with one aspect of behavior as to those primarily concerned with total behavior. It was found that weekly ratings picked up behavioral variations (both cyclical and progressive-regressive) better than shorter or longer intervals. Although rater selection was random in the present study, it is felt that better results could be obtained by asking the ward charge to select for rating those aides who are good at paper work, thus getting the more intelligent aides. (The fewer invalid ratings made by the total push group of raters were felt to be due to the fact that they are among the best aides in the hospital, and are in closer contact with the patients). Halo effect is reduced by having separate sheets for rating each time so that raters may not see their previous ratings of the patients.

Uses. The possible uses are legion, and only a few will be mentioned here. The most obvious is for evaluation of patients when other psychometric tools cannot be used. In research it can be used for evaluating therapy, comparison of different hospital populations, picking up differences between various clinical types of patients, correct ward placement of patients, catching *spontaneous* remissions, etc.

Summary

The L-M Fergus Falls Behavior Rating Scale provides an inexpensive, convenient, fairly objective and reliable method of rating the behavior of patients in mental hospitals. Inexperienced raters can use it with facility, obtaining agreements of 89 to 97 percent. It measures eleven aspects of behavior, and is sensitive enough to pick up differences between patients and changes in one patient over a length of time. It can be used when other psychometric techniques are ineffectual (when patient is mute, agitated, etc.), and in research involving evaluations of behavior.

References

1. Cohen, L. H., Malmo, R. B., and Thale, T. Measurements of chronic psychotic overactivity by the Norwich Rating Scale. *J. gen. Psychol.*, 1944, 30, 65-74.
2. Janis, Irvin L. Psychologic effects of electric convulsive treatments. (III. Changes in affective disturbances). *J. Nerv. Ment. Dis.*, 1950, 111, 469-489.
3. Malamud, W., Hoaglund, H., and Kaufman, I. C. A new psychiatric rating scale. *Psychosom. Med.*, 1946, 8, 243-245.
4. Malamud, W., and Pincus, G. Changes in the electroencephalogram and in the excretion of 17 ketosteroids accompanying electro-shock therapy of agitated depression. *Psychosom. Med.*, 1946, 8, 246-251.
5. Malamud, William, and Sands, Sidney L. A revision of the Psychiatric Rating Scale. *Amer. J. Psychiat.*, 1947, 104, 231-237.
6. Sands, Sidney L., and Malamud, William. A rating scale analysis of the clinical effects of lobotomy. *Amer. J. Psychiat.*, 1949, 105, 760-766.
7. Wittman, Phyllis. The Elgin check list of fundamental psychotic behavior reactions. *Amer. Psychologist*, 1948, 3, 280.

AN OBJECTIVE METHOD OF EVALUATING MENTAL STATUS[1]

JOHN T. ROWELL, R. N., M. A.

Florida State University

Introduction

In this study a behavioristic approach to the measurement of psychotic behavior in a ward environment has been attempted. This was initiated on the assumption that objective behavior gives one of the most important clues to the adjustment of an individual in a social situation, whether it be within a restricted environment such as we find in hospitals or the freer atmosphere outside.

Guthrie (2) contends that "practical and sound description of personality traits must be made in terms of learned skills and adjustments rather than in terms of general traits like 'introvert', 'ascendant', or the like". He believes that "people can sometimes be trusted to do the expected thing in a large variety of situations", and that "what a man is doing at any time is a most important determiner of what he will do next". Cameron (1) suggests that in many psychotics, especially the schizophrenics, there is a maladjustment in the ability of the individual to take roles in a social situation, usually arising from poor parental support and frustration at an early age. He speaks of the schizophrenic as a person "who has never acquired the degree of social skills he needs for shifting his perspectives through taking successive culturally determined roles when he is under stress".

The purpose of this study was to ascertain whether or not a graphic rating device could be constructed which would meaningfully portray the behavior of psychotic patients in a ward environment. In an effort to arrive at this aim, two tentative hypotheses were set up: (a) that a statistical method could be devised which would yield a score indicative of the position of any given individual relative to society, and (b) that these scores when plotted graphically would have meaning and some predictive value.

Procedure and Techniques

Twenty-four "behaviors" were selected after careful examination of standard texts on psychiatry for descriptive terminology. No attempt was made to divide characteristics such as "delusions" as to types. These items were then used to construct item scales with five discriminators each, describing various behavior intensities ordinarily seen among mental patients. Once this was completed, the items were presented to twenty-two psychiatrically trained nurses for comments and suggestions in an attempt to remove ambiguity. Four items were eliminated and the final scale as used consisted of twenty items called the *Psychiatric Behavior Scale* which is reproduced in Fig. 1.

A period of instruction was given to the nursing staff prior to the actual rating of the patients. The boundaries of each item were discussed and the mechanics of recording given. An instruction sheet was placed with each set of scales so that the nurse might refer to it if necessary. Day nurses were instructed to evaluate the patient in retrospect for an eight hour period in blue ink and the evening nurses to do the same in red ink. Night nurses did not check the *Scale* since there was no attempt made to describe behavior during hours of sleep. In addition to the two daily ratings, the nurse recorded any unusual conversation or behavior of the patient as well as any treatment that the patient may have received such as electroshock or insulin therapy.

---

[1]Data for this study were compiled in 1949 at the Neuropsychiatric Institute, University of Michigan, Ann Arbor, Michigan. The writer is greatly indebted to the staff and personnel who so willingly gave of their time, and to various members of the Florida State University, Tallahassee, Florida, for a critical reading of the text.

Figure 1. Psychiatric Behavior Scale

NAME..... DATE.....

Rate each item at the mark, not between marks.

1. WAS PATIENT *PREOCCUPIED?*

Seldom | Part of the time | About half the time | Most of the time | Completely

2. DID PATIENT *HALLUCINATE?*

Absent | Appeared to | Occasionally | Frequently | Freely

3. WAS PATIENT *DELUSIONAL?*

Absent | Appeared to be | Occasionally expressed | Frequently but unorganized | Highly and organized

4. DID PATIENT SHOW *INSIGHT* INTO ILLNESS?

Good | Good at times | Fair at times | Doubtful | None

5. DID PATIENT *APPRAISE HIS ACTIONS* IN REGARD TO THE FEELINGS OF OTHERS OR TO THE VALUE OF PROPERTY?

Always | Most always | Oscillated | Almost never | Never

6. PATIENT'S *AFFECT* WAS?

Evenly expressed | Usually evenly expressed | Oscillated | Flattened | Severely flattened

7. STATE OF PATIENT'S *ORIENTATION* WAS?

Well oriented | Seldom confused | Oscillated | Confused most of time | Highly confused

8. PATIENT'S *MOOD* WAS?

Even | Mildly elated | Elated | Euphoric

Oscillated

Mildly depressed | Depressed | Melancholic

9. DID PATIENT *SOCIALIZE?*

Almost constantly | Frequently | Occasionally with difficulty | Seldom | Never

10. PATIENT'S *MOTOR ACTIVITY* WAS?

Even | Slightly retarded | Retarded | Extremely retarded

Oscillated

Slightly hyperactive | Hyperactive | Extremely hyperactive

11. DID PATIENT SEEM *AGITATED?*

No | Slightly agitated | Oscillated | Agitated | Highly agitated

12. DID PATIENT ATTEND *ACTIVITIES?*

Attended all | Attended most | Indifferent | Refused most | Refused all

13. *APPETITE?*

Ate all | Ate most | Ate fairly well | Refused most | Refused all

14. WHAT WAS PATIENT'S *ATTITUDE TOWARD PERSONNEL?*

Friendly | Accepting | Indifferent | Avoiding | Fearful or menacing

15. PATIENT'S *PERSONAL APPEARANCE* WAS?

Well groomed	Neat	Occasional interest in appearance	Slovenly	Soiled

16. PATIENT'S GENERAL *BODY TONE* WAS?

Relaxed	Slightly tense	Oscillated	Tense	Extremely tense

17. DID PATIENT SHOW *BLOCKING OF SPEECH?*

Absent	Appeared to	Slightly	Moderately	Severely

18. WAS PATIENT NEGATIVISTIC?

No	Occasionally	Frequently but unsystematized	Frequently and Systematized	Physically Resistive

19. DID PATIENT EXPRESS *SOMATIC COMPLAINTS?*

Absent	Seldom	Occasionally	Frequently	Persistantly

20. DID PATIENT SHOW *"FLIGHT OF IDEAS"*?

Absent	Occasionally	Oscillated	Frequently	Consistently

REMARKS: (Any unusual action or conversation of patient.)

TOTAL SCORE..... NURSE.....

Forty-four patients, twenty-three male and twenty-one female, were used for the study. Patients were taken as admitted to the Neuropsychiatric Institute of the University of Michigan at Ann Arbor. The *Scale* was administered for a period of thirty-one days. Fifteen patients were admitted during the study, twenty were in residence throughout, and nine were discharged while the collecting of data was in progress.

The *Psychiatric Behavior Scale* with its twenty items on a five point continuum yielded scores from twenty to one hundred. Each point on the continuum was arbitrarily considered to be of equally increasing weight and no attempt was made to allow for more than five discriminators. Following the collection of data the day and evening scores were averaged to obtain a total score for the patient's behavior for the day. This score was then plotted graphically along with any treatments or special comments of interpretative value (see Fig. 2).

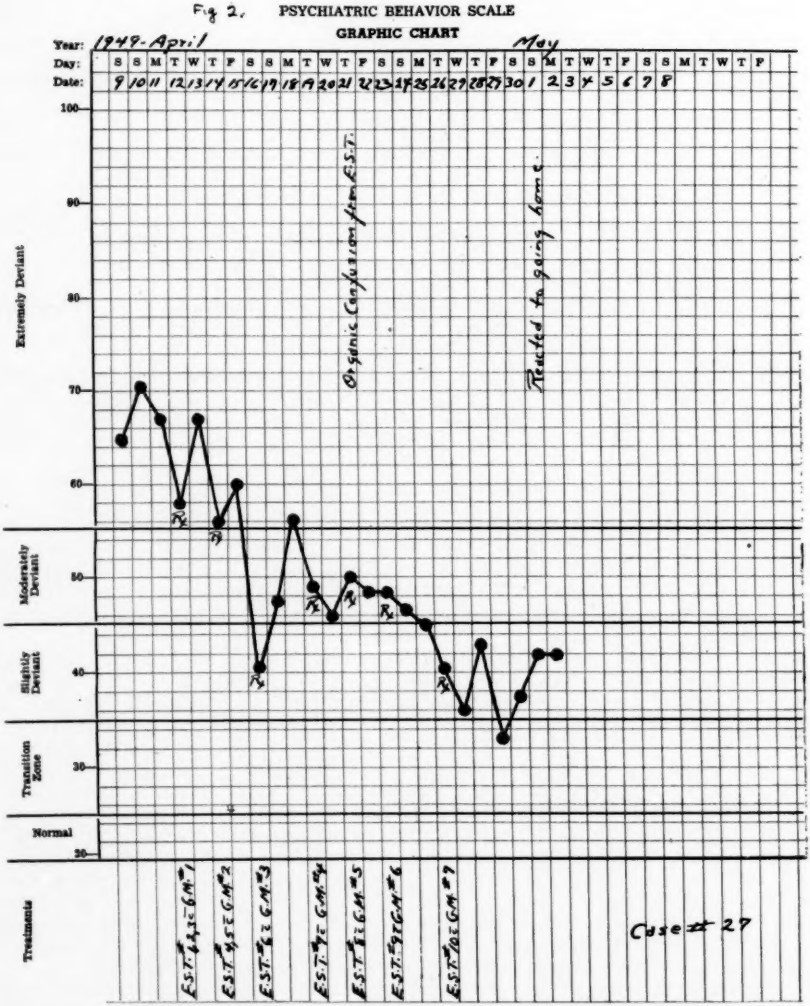
Several correlations were obtained during the period of administration from paired estimates. A form of test-retest was given each nurse to determine consistency of ratings. In this a second group of rating sheets were presented immediately upon completion of the first group of ratings. The nurse had no prior knowledge that she would be required to re-evaluate her previous decisions. This was called the "test-retest" estimates. Even though it was felt that some immediate memory recall was involved, the nurses reported that they referred back to the patient's condition when rerating the scales rather than trying to recall their previous choices.

Reliability was also determined by having two psychiatrically trained nurses rate independently the same patients for the same period of time. Eight nurses were used to obtain sixty-two paired ratings. From this data a reliability coefficient was computed. This was referred to as the "nurse-nurse" estimates.

Two methods were used to obtain validity for the scale. Arrangements were made whereby the eight psychiatrists who were immediately directing the treatment of the patients studied, checked the *Scale* weekly in retrospect as well as ranked their patients.

The ratings made by the psychiatrist for each patient were paired with the mean rating of that same patient by the nurses on the ward and a validity coefficient was computed. There were ninety-four pairs used to compute this coefficient, these in-

Fig 2. PSYCHIATRIC BEHAVIOR SCALE


GRAPHIC CHART
Year: 1949-April
May
Day: S S M T W T F S S M T W T F S S M T W T F S S M T W T F S S M T W T F
Date: 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 1 2 3 4 5 6 7 8
100
90
80
70
60
50
40
30
20
Extremely Deviant
Moderately Deviant
Slightly Deviant
Transition Zone
Normal
Treatments
Organic Confusion from E.S.T.
Reacted to going home.
E.S.T. 42 = 6 M.V. 1
E.S.T. 45 = 6 M.V. 2
E.S.T. 52 = 6 M.V. 3
E.S.T. 72 = 6 M.V. 4
E.S.T. 81 = 6 M.V. 5
E.S.T. 91 = 6 M.V. 6
E.S.T. 102 = 6 M.V. 7
Case # 29

cluded the ratings of fourteen psychiatric nurses and eight psychiatrists. These paired estimates were referred to as the "nurse-psychiatrist" estimates.

In ranking their patients, the psychiatrists were also asked to arrange them in rank order from the patient they considered nearest to society, i.e., the one who would be observed as least different if in society, to the one farthest away from what society considers a *normal* or *average* person. These rankings were then compared with the rank-order determined by scores obtained when the psychiatrist rated the *Scale*. This was an effort to determine if the scale was actually showing the amount of deviant behavior of the patient as implied by the total scores.

Analysis of Data

Ratings were completed within thirty-one days on the forty-four cases. An analysis of the data yielded encouraging results. The "test-retest" estimates, using seventy-seven pairs, yielded a Pearson product-moment r of .95. The "nurse-nurse" estimates, using sixty-two pairs, yielded an r of .85. Both of these were found to be highly significant. The "nurse-psychiatrist" judgments, involving ninety-four pairs, were found to correlate .78. Using the Fisher's *Z* function for given values of r, we can be confident at the one percent level that the true r for the "nurse-psychiatrist" distribution of judgments will fall between .65 and .87.

The psychiatrists' ranking of patients in relation to the score ranking by the scale yielded seventy-one complete agreements out of eighty-eight rankings for a total agreement of eighty-one percent. Figure 2 graphically portrays a forty-four year old married female who was admitted with a provisional diagnosis of *Menopausal Psychosis.* She was later diagnosed by the staff as *Involutional Melancholia with Depression.* On admission she was preoccupied most of the time and frequently expressed somatic delusions which were of an organized nature. She showed no insight into her illness and seldom appraised her actions in regards to the feelings of others. Her affect was severely flattened and she was very depressed most of the time. The patient made no attempt to socialize with others on the ward and showed extreme motor retardation. She refused all activities and ate very poorly. She was indifferent, avoiding of personnel, and tense on approach. Hallucinations were absent and no flight of ideas was shown. Increasing agitation appeared following admission characterized by pacing of the floor, crying, and a general increasing of motor activity.

The severity of her behavior was clearly demonstrated by the position of the first recordings on the *Scale.* Her total score as shown in Figure 2 was sixty-five which located the patient in the *Extremely Deviant* classification. On the fourth day following the first recording, she received her first electrical stimulation and as the graph shows, she dropped nine points in the direction of *Normal.* The following day she showed a climb back to the position held before the first treatment. As will be noted, the patient continued to show further drops toward the *Normal* classification with each electrical stimulation in a step-ladder fashion until the fifth treatment when a rise in the amount of deviant behavior was shown. The psychiatrist in charge of the case confirmed organic confusion at this point brought on by the electrical treatments. The sixth stimulation showed a neutralizing effect obtained when overall improvement was antagonistic to organic confusion. This was confirmed by a further drop in the total score on subsequent days.

Conclusions

It is felt that this preliminary study demonstrates that psychotic behavior can be recorded numerically and that this numerical recording yields information which may facilitate a better understanding of the patient. As demonstrated above, a plotting of the scores graphically will allow for interpretation and prediction as well as allow for an accurate historical review of the patient's illness.

There is also an indication that valuable data would result from a longitudinal administration of the *Psychiatric Behavior Scale* to a large population of patients. A comparison of graphs from this study points to many similarities between patients of different psychiatric classifications during the course of their illnesses. A longitudinal recording would allow for better comparison and interpretation. Studies are now in progress to determine whether or not psychiatric "types" show characteristic behavior scores. A report of these studies will be made at a later date.

Bibliography

1. Cameron, Norman. *The Psychology of Behavior Disorders.* New York: Houghton Mifflin Co., 1947. Pp 485-86.
2. Guthrie, Edwin R. Personality in Terms of Associative Learning, in *Personality and the Behavior Disorders* (ed. by J. McV. Hunt). New York: The Ronald Press Co., 1944.

AN ANALYSIS OF THE ELGIN PROGNOSTIC SCALE*

MAURICE LORR, PHYLLIS WITTMAN, AND W. SCHANBERGER

Elgin State Hospital

Introduction

Clinicians are constantly seeking factors useful for the prediction of recovery in the functional psychoses. Prognostic factors are much needed in the selection of patients for treatment procedures such as lobotomy or insulin shock. They are equally important to research workers for the proper matching of patients with control subjects in experimental studies. The Elgin Prognostic Scale has much to offer towards satisfying these needs.

The Elgin Scale is a rating schedule designed to predict recovery in schizophrenics. It consists of 20 rating scales weighted according to prognostic importance; favorable factors are arbitrarily assigned negative weights and unfavorable factors are assigned positive weights. The prognostic score is the algebraic sum of the weighted measures. The variables are in large part based upon premorbid social history secured from the patient's relatives.

These scales were constructed following a review by Wittman [6] of over fifty studies concerned with prognostic factors in schizophrenia. The initial validation study included 343 cases diagnosed by the medical staff at the Elgin State Hospital as schizophrenic and placed on various types of shock treatment. Staff judgments of prognosis and prognostic scale ratings were secured independently before treatment on each patient. The scales were effective in predicting immediate outcome of therapy and proved to be much superior to staff estimates of prognosis. In a follow-up study [7] on 804 schizophrenic and 156 manic-depressive patients, comparisons were again made between the prognostic scale ratings and staff judgments of prognosis relative to clinical status following treatment. Again the ratings were found to predict outcome of therapy with a greater accuracy than staff judgments.

The purpose of this present study is to isolate and identify by factor analysis the independent factors which will most parsimoniously describe the interrelation of

Table 1. Centroid and Rotated Factor Loadings and Communalities

	Centroid Factor Matrix					Rotated Factor Matrix		
No.	I	II	III	h^2	No.	X	Y	Z
1	79	08	37	77	1	86	17	-04
2	85	17	26	81	2	85	32	-02
3	84	13	41	90	3	93	19	-08
4	65	-28	26	57	4	65	03	24
5	70	-12	-16	53	5	47	43	40
6	92	-07	13	87	6	81	36	26
7	96	-03	18	95	7	87	35	21
8	66	10	-23	50	8	41	56	27
10	48	-68	-36	82	10	12	23	84
11	30	-75	-32	76	11	-01	08	84
12	68	14	25	54	12	70	23	-04
15	62	-15	42	58	15	73	-05	04
16	38	40	-58	63	16	01	79	17
17	26	46	-44	47	17	00	65	01
18	65	20	-22	51	18	41	59	19
19	73	35	-10	67	19	57	60	03
20	74	20	11	59	20	68	38	01

*Sponsored by the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

20 scales in the rating schedule. Such an analysis might be expected to: (a) provide a better understanding of what determines recovery; (b) increase the clarity with which each scale can be constructed; and (c) throw light on the validity of certain hypotheses regarding schizophrenia.

Method and Results

Two hundred successive admissions at the Elgin State Hospital were rated on the Elgin Prognostic Scale by one of the authors and a psychiatrist. Tetrachoric correlation coefficients between 17 of the 20 scale variables were computed. The resulting correlation matrix was factored by the centroid method. The three factors obtained were then rotated to oblique simple structure by the method of extended factors. The centroid and rotated factor loadings are presented in Table 1, page 260. Decimal points are omitted to save space.

On the first factor X, 13 of the variables have factor loadings or weights of .35 or more. These variables, labeled by the positive or unfavorable end of each scale, are given below.

3 Defects of Interest	.93
7 Insidious Onset	.87
1 Shut-In Personality	.86
2 Schizothymic Personality	.85
6 Limited Range of Interests	.81
15 Constitutional Bias	.73
12 Low Energy Tone	.70
20 Asthenic Build	.68
4 Lack of Heterosexual Contacts	.65
19 Absence of Atypical Symptoms	.57
5 Inadequate Affect	.47
8 Long Duration of Psychoses	.41
18 Hebephrenic Symptoms	.41

Variables high on this factor appear to have in common the indications of schizoid withdrawal which is typically insidious. A constitutional predisposition towards a low level of energy and drive appear to be prominent in this factor. A patient having a high score on this factor would be described as a "shut-in" personality, withdrawn, seclusive, with few or no interests or friends. From early childhood he has been regarded as queer or odd. Not only has he had few friends but he probably has had virtually no social contacts with members of the opposite sex.

Schizoid withdrawal as represented by factor X is commonly regarded as a frequent precursor of schizophrenia. Schizoid withdrawal is described by Jenkins [2] in a recent paper as a "relative withdrawal of attention and interest from outer environment, particularly a withdrawal from empathic contact with the human environment . . ." Freud [1] characterized this element as follows: "In schizophrenia, on the other hand, we have been obliged to assume that after the process of repression the withdrawn libido does not seek a new object, but retreats into the ego; that is to say, here the object-cathexes are given up and a primitive object-less condition of narcissism is re-established. The incapacity of these patients for transference—so far as the process of the disease extends—their consequent inaccessibility to therapeutic efforts, the repudiation of the outer world characteristic of them, the manifestations of hypercathexis of their ego, the final outcome in complete apathy—all these clinical features seem to accord excellently with the assumption that object-cathexes are relinquished."

The second prognostic factor is denoted Y. The principal factor loadings on this reference axis are as follow:

16 Ideas of Influence	.79
17 Bizarre Delusions	.65
19 Absence of Atypical symptoms	.60

18 Hebephrenic Symptoms	.59
8 Long Duration of Psychoses	.56
5 Inadequate Affect	.43
20 Asthenic Build	.38
7 Insidious Onset	.36
6 Limited Range of Interests	.36

Ideas of influence (16) and bizarre delusions (concerning bodily organs and process) have in common the distortion of reality subsequent to the cessation of reality-testing. Hebephrenic symptoms (18) are indicative of such distortion and of the divorce between ideation and affect resulting from extreme personality disorganization. Of interest also are the factor loadings of scales (8) and (7). The more gradual the onset and the longer the duration of the psychosis, the greater the likelihood of personality disorganization and loss of contact with reality.

Jenkins [2] recently suggested that schizophrenia is a progressive maladaptation occasioned by frustration beyond the tolerance of the individual. Interpreted in terms of Jenkins' concept of the schizophrenic process, this factor is descriptive of the change which occurs in the individual when he passes his threshold of tolerance for frustration and begins to react in nonadaptive fixed ways.

While the first factor reflected the gradual process of schizoid withdrawal, the second factor reflects the less gradual process of schizophrenic distortion. Concomitant with progressive schizophrenic withdrawal there is a gradual reduction in reality-testing. As a result, reality distortion following the psychotic break is likely to be great. Factor Y appears to correspond to the factor identified by Moore [4] in his early study as the "Deluded and Hallucinatory Syndrome."

The scales having significant loadings in the third prognostic factor Z are as folows:

11 Stubborn Traits	.84
10 Careless Indifference	.84
5 Inadequate Affect	.40

The positive end of scales (10) and (11) describes the rigid, stubborn, opinionated individual who refuses to alter his accustomed patterns or ways. His disregard of social requirements is reflected in his language, appearance and habits. These scales suggest a factor of inadaptability or rigidity. In further studies it should be possible to construct new scales to measure this factor and to determine whether it can be more usefully defined for prognostic purposes.

Since the three centroid factors were rotated to an oblique structure not all of the reference axes are completely independent. The factors representing reality distortion and rigidity have a slight negative correlation. Schizoid withdrawal and rigidity show a small positive correlation. Schizoid withdrawal and reality distortion as represented by factors X and Y are statistically independent. The transformation matrix containing the direction cosines of the reference vectors are given below.

	X	Y	Z
I	.810	.521	.308
II	.092	.416	−.735
III	.578	−.744	−.603

Summary

Two hundred hospitalized psychotics were rated on the Elgin Prognostic Scale. The scale ratings were intercorrelated and the matrix of correlations factored by the centroid method. The three factors obtained may be interpreted as (1) a factor of schizoid withdrawal, (2) a factor of schizophrenic reality distortion, and (3) a less well-defined factor of personality rigidity or inadaptability. The first two factors seem identical with schizoid withdrawal and schizophrenic disorganization which are basic elements of the hypothesis of schizophrenia recently presented by Jenkins, and to this extent our findings support this hypothesis.

Bibliography

1. Freud, S. The unconscious. *Collected Papers.* London; Hogarth Press, 1925, Vol. IV, 128-129.
2. Jenkins, R. L. Nature of the Schizophrenic process. A working hypothesis for therapy. *Arch. neur. Psychiat.*, 1950, *64*, 243-262.
3. Meyer, A. *The commonsense psychiatry of Dr. Adolph Meyer.* New York; McGraw-Hill, 1948.
4. Moore, T. V. The essential psychoses and their fundamental syndromes. *Studies in psychol. and psychiat. from the Cath. Univ. of Amer.*, 1933, III.
5. Thurstone, L. L. *Multiple factor analysis.* A development and expansion of the vectors of mind. Chicago: University of Chicago Press, 1949.
6. Wittman, P. A scale for measuring prognosis in schizophrenic patients. *Elgin papers* (Elgin State Hospital, Illinois), 1941, *4*, 20-33.
7. Wittman, P. and Sternberg, L. Follow-up of an objective evaluation of prognosis in dementia praecox and manic-depressive psychoses. *Elgin papers*, 1944, *5*, 216-227.

THE USE OF TOLSEROL (MYANESIN) IN PSYCHOLOGICAL TESTING*

MARGARET MERCER, PH.D. AND ARTHUR O. HECKER, M. D.

Veterans Administration Hospital
Coatesville, Pennsylvania

Introduction

It is the purpose of this paper to discuss the possible uses of Tolserol in psychological testing in the neuropsychiatric hospital. Myanesin (Tolserol) is a drug having a transitory effect upon the midbrain in appropriate dosage. Clinically, it has been demonstrated that this drug will temporarily relieve anxiety. Studies made at this hospital to investigate the clinical effects of Tolserol (myanesin) suggest that this drug may be useful in studying the effects of psychosis upon intellectual functioning in situations where it is desirable to know whether lowered efficiency results from anxiety and blocking, or reflects psychotic deterioration unlikely to change. The psychological aspects of the two exploratory studies previously made will be briefly summarized.

The first study[1] investigated the effects of Tolserol on twenty-six schizophrenic patients. Duration of the illness varied from two to twelve years. The majority of the patients had psychotherapy, electro-shock therapy, deep insulin shock therapy and/or combined deep insulin and electroshock therapy. The group in its entirety could be considered as having well established mental illness which was not amenable to available forms of therapy. All patients received one gram of Tolserol every two hours five times per day for a three week period. Psychological testing consisted of the verbal portion of the Wechsler-Bellevue Scale and a Rorschach examination given one week before therapy was initiated and repeated immediately following therapy. At the same time, a patient carrying the diagnosis of chronic anxiety state received similar treatment with Tolserol and was similarly tested. The results of the psychological tests given to the schizophrenic patients reflected only the slight changes which fell within the ordinary range of re-test variation. However the patient with the diagnosis of anxiety state showed an 18 point improvement in the verbal IQ. Since Tolserol is known to be rapidly detoxified in

*Sponsored by the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are a result of their study and do not necessarily reflect the opinion or policy of the Veterans Administration.

the body there was no reason to expect that the patients would show any significant differences in test results when tested after drug therapy was completed. However, the neurotic patient did so. It was assumed that this patient had probably profited from the tension-free periods which had resulted from the Tolserol to become generally less tense. Decreased tension manifested itself in increased ability to attend and concentrate as well as in more free and complete response to questions on the comprehension and similarities sub-tests. Some of the schizophrenic patients had shown on the wards minor changes in behavior characterized by less tension and greater communicativeness. Because of the rapid detoxification of the drug only transitory changes would be expected.

A second study[2] of clinical experiences with Tolserol was made on twenty-six patients in various diagnostic categories selected for investigation because emotional factors were dominating their reactions in a fashion which blocked effective therapeutic rapport or general hospital adjustment. Responses to Tolserol fell into five general types:

1. Release of blocking to recall and revelation of painful experiences with or without concomitant emotional release. In the former instance, a level of violent abreaction was obtained in one case.

2. Control of vivid emotional reaction (rage) in manic type of release.

3. Relief of overt anxiety with more logical sequential alignment and expression of mental content.

4. A combination of the above three responses.

5. No response (this group was made up of patients in whom the psychosis had become paramount so that events of the external world no longer needed to be dealt with and anxiety was no longer needed.)

Blind analyses of these patient's psychological examinations were found to yield remarkably accurate predictions of the form which reaction to Tolserol had actually taken.

Procedure

As a result of these two studies it was decided to test some patients while they were directly under the influence of Tolserol. Thirteen patients were selected whose verbal Bellevue-Wechsler scores were characterized by marked scatter with vocabulary and general information higher than the other sub-test scores. These patients were similar to those studied in the first paper previously summarized. All were patients who carried a diagnosis of schizophrenia. Mental illness was well established. The patients had proved unresponsive to psychotherapy and various forms of shock therapy.

The patients were given two and one-half grams of elixir of Tolserol by oral administration. The drug required about 20 minutes to take effect and did not wear off for about forty minutes. It was, of course, necessary that the testing time be carefully scheduled by the psychologist testing the patient and the physician administering the drug. It was the experience in previous studies that spontaneous nystagmus was the criterion for noting when the dosage was effective and that two and one-half grams was the amount most commonly needed to produce this condition. In this study the dosage was kept constant to maintain standard conditions during the experiment. Tests given under Tolserol were the comprehension, digit span, arithmetic and similarities sub-tests of the Wechsler-Bellevue Scale and the free associations to the Rorschach cards.

Results

The results of Wechsler-Bellevue testing with and without Tolserol are shown in Table 1.

TABLE 1. WECHSLER-BELLEVUE SUBTEST SCORES ON THIRTEEN SCHIZOPHRENIC PATIENTS WHEN TESTED (A) WITH AND (B) WITHOUT TOLSEROL

	Comp.	Digit Span	Arith.	Sim.	Total	Diff.
A	9	6	4	10	29	
B	7	4	0	9	20	−9
A	4	6	7	3	20	
B	4	7	3	2	16	−4
A	11	11	13	12	47	
B	12	11	13	10	46	−1
A	5	4	3	10	22	
B	5	4	3	10	22	0
A	5	4	6	7	22	
B	5	8	4	6	23	+1
A	6	9	3	4	22	
B	9	7	4	3	23	+1
A	9	11	13	12	45	
B	12	11	12	12	47	+2
A	5	10	9	13	37	
B	6	13	13	12	44	+7
A	9	7	4	10	30	
B	11	7	9	11	38	+8
A	5	4	3	3	15	
B	7	4	7	6	24	+9
A	9	14	10	10	43	
B	12	17	13	12	54	+11
A	10	4	12	14	40	
B	15	6	16	15	51	+11
A	1	0	0	1	2	
B	7	3	0	8	18	+16

It will be seen that these thirteen schizophrenic patients differed markedly in their response to test situations with and without Tolserol. In one case, relaxation of tension resulted in significantly poorer response. The patient verbalized feelings of greater relaxation and said "I will be able to rest and relax." He was obviously less tremulous and anxious while Tolserol was effective. This appears to be an example of the type of personality in which anxiety acts as a motivating force to more effective performance. In six cases the amount of variation was of no significance. In six other cases there was a degree of improvement which would cause an increase in the verbal IQ greater than the limits of probable error of the Wechsler-Bellevue Scale would lead one to expect. The data summarized in Table 1 shows sufficient changes in intellectual functioning reflected in the Wechsler-Bellevue sub-tests as to justify further investigation of the hypothesis that testing under the influence of Tolserol may provide a means of determining whether the patient retains the capacity for change or whether the psychotic process has advanced to such a degree that irreversible damage rather than loss of efficiency has taken place.

In addition to the Wechsler-Bellevue material the patients were also given the Rorschach while under the influence of Tolserol. Because of the relatively short period in which the drug operates, only the free association period of testing was carried out. No inquiry or testing of the limits were done. As the Rorschach reflects

the basic personality structure it would not be expected that temporary alleviation of tension would produce significant change. This proved to be the case. The associations produced under Tolserol were essentially the same as those produced without it.

Lowered efficiency of intellectual functioning is often observed in cases where there has been organic damage to the central nervous system. It is often naively assumed that all lowered functioning in organic cases is due to the organic condition, that it is permanent, and that nothing can be done to aid the patient in re-establishing a more effective level of function. In the course of our exploratory studies we have tested four organic patients and have found that two of these cases show no change in Bellevue-Wechsler scores under Tolserol and that two showed significant improvement.

Conclusions

Exploratory study of psychological testing under the influence of Tolserol suggests that in the neuropsychiatric hospital this may be a useful means of distinguishing between loss of efficiency due to the impact of existing anxiety and that resulting from the inroads of the disease process. It may also provide a method of following the course of the disease process in a more precise fashion than clinical observation permits.

Bibliography

1. Dinenburg, Samuel and Hecker, Arthur O.: Oral Myanesin (Tolserol) in Schizophrenia. *Dis. Nerv. Sys.*, 1950, 11, 2.
2. Hecker, Arthur O., Mercer, Margaret, and Griffin, Mark A., Jr.: Clinical Experiences with Tolserol. In preparation.
3. Dixon, Henry H., Dickel, Herman A., Coin, Robert A. and Hauger, Gerhard B.: Clinical observations on Tolserol in handling anxiety tension states. *Am. J. Med. Sci.*, 1950, 220, 23-29.

A COMPARISON OF THE PERFORMANCE OF A RANDOMLY SELECTED COLLEGE POPULATION ON THE MMPI AND THE P-S EXPERIENCE BLANK

PAUL W. CAUFFIEL AND WILLIAM U. SNYDER

The Pennsylvania State College

Introduction

Considerable speculation arises in clinical situations when attempts are made to evaluate individuals against norms provided with the various tests. Are these norms valid for local populations? If an individual shows maladjustment on one test, would he escape screening on another? Are the so-called "validating" scales of many tests of value?

The present study reflects these questions to a large extent, supplies evidence of some homogeniety in random college populations on two tests, and indicates in this instance that disturbed individuals may be screened by both measuring devices. It also suggests that "validating" scales may be useful indicators of maladjustment by themselves.

The experiment was designed to study the linear relationship between the subscores of the Minnesota Multiphasic Personality Inventory (MMPI) and of the Psycho-Somatic Experience Blank (PSEB), and to compare the responses of various groups on the MMPI and the PSEB.

The PSEB is less well known than the MMPI. It consists of two forty-six-item scales; one related to physiological symptoms and the other to psychological symptoms. The total score is derived by a combination of the scores of the two scales. Individuals can be compared with both normal and neurotic male and female populations[5, 6].

Related Studies

Brown[1] found that several college populations, although somewhat similar in profile on the MMPI, differed significantly. He suggests that it may be advantageous to develop local norms for personality tests as well as for tests of ability. He found that all college groups in his study had slightly elevated Hy scores when considering the "neurotic triad" components of the MMPI. Fry's population[2] showed a similar characteristic.

The K scale of the MMPI was derived by studying the responses of abnormals producing normal profiles, assuming that the resulting profile was normal because of defensive attitudes in the patients' responses. It was found that university students have a relatively higher mean K score than others within the same age range. This difference among college students is thought to be chiefly a function of socio-economic factors[9]. The usual interpretation of a high K attributes it to defensiveness and a low one to unusual frankness or self-criticism[8].

Procedure

Two hundred ninety students of The Pennsylvania State College were selected from the *Student-Faculty Directory* by the page-sampling method. Names of male students without cards in the Registrar's file were omitted and the following name substituted. An exception was made for women students, wherein the next *woman's* name was substituted in an effort to keep the sex ratio constant. Two hundred one subjects completed the study. One hundred fifty were males ranging in age from 18 to 39 with a mean of 23.07 $\pm$ 2.97. The fifty-one women tested range in age from 18 to 34 with a mean age of 21.37 years $\pm$ 2.76. Fourteen testing sessions were conducted with groups of approximately fifteen subjects each. The order of test administration was alternated with each group to avoid possible biasing effects.

The subjects were divided into four groups; (1) total group of 201, (2) 150 men, (3) 51 women, and (4) the deviant group, composed of the twenty-five scoring lowest on the PSEB total scale. The four groups were compared with each other by calculating the standard error of the difference between means for uncorrelated samples, and then calculating the critical ratio of the difference between means. Pearson product-moment correlations between the subtests of the MMPI and the PSEB were determined (table 1). Means and standard deviations were calculated (table 2). for all scales for the four groups.

TABLE 1. LINEAR RELATIONSHIP BETWEEN THE SUBTESTS OF THE MMPI AND THE PSEB SCALES (N 201)*

MMPI	PSEB Physiological	PSEB Psychological
Hs	−.535	−.378
D	−.395	−.544
Hy	−.077	.054
Pd	−.187	−.293
Mf	−.130	−.148
Pa	−.208	−.120
Pt	−.617	−.702
Sc	−.521	−.635
Ma	−.132	−.217
K	.446	.576
F	−.344	−.391

*A negative correlation is, in this case, indicative of a positive correlation, because of the nature of the scoring of the two tests.

TABLE 2. MEAN SCORES AND STANDARD DEVIATIONS FOR THE MMPI AND THE PSEB OBTAINED BY A RANDOMLY CHOSEN POPULATION

Test Scale	Total M	Total σ	Men M	Men σ	Women M	Women σ	Deviant M	Deviant σ
Phys.	119.77	40.72	119.41	41.49	120.84	38.33	50.92	34.41
Psych.	125.51	58.57	128.58	59.17	116.49	56.65	17.88	47.23
Total	243.87	89.29	245.07	83.57	240.31	104.17	68.80	62.08
Hs.	3.63	3.02	3.76	3.11	3.25	2.71	7.00	4.21
D	17.77	4.26	17.80	4.37	17.69	3.89	23.32	3.87
Hy	20.67	3.66	20.65	3.93	20.73	2.75	21.64	3.97
Pd	14.63	3.94	14.83	3.97	14.06	3.72	17.56	4.79
Mf			25.21	5.01	36.86	5.33		
Pa	8.77	2.60	8.65	2.62	9.10	2.12	9.88	2.96
Pt	9.14	6.64	9.18	6.95	9.04	5.62	18.84	6.15
Sc	8.61	6.56	8.82	6.75	8.00	5.67	17.32	8.84
Ma	16.98	3.98	17.24	4.11	16.22	3.45	19.36	5.21
K	15.84	4.45	15.69	4.59	16.27	4.03	11.00	3.45
F	3.86	3.31	4.05	3.49	3.29	2.64	7.00	3.51
	N 201		N 150		N 51		N 25	

RESULTS

Table 1 shows the linear relationships between the two scales of the PSEB and the MMPI. Low scores on the PSEB indicate maladjustment while the converse is true for the MMPI; all negative correlations, therefore, indicate direct linear relationships.

Pt correlates higher with both scales of the PSEB than all other scales of the MMPI. This relationship may be explained by the similarity of items on these three scales. *Hs* correlates substantially with the physiological scale of the PSEB and not quite as heavily with the psychological scale. This is to be expected, since hypo-

chondriacs would usually be more concerned with physical symptoms than others. McKinley and Hathaway [7] found that the presence of physical symptoms only slightly altered the personality pattern in the direction of hypochondriasis. *Sc* correlates substantially with both scales of the PSEB. Hathaway and McKinley [4] found the correlation for normal subjects between *Pt* and *Sc* to be .84. This relationship may be due to the overlap of the items used on these scales. The *D* scale shows substantial relationship to the psychological scale of the PSEB and a less pronounced relationship to the physiological scale. It is difficult to explain the negligible correlations between *Hy* and both scales of the PSEB unless the subtle and obvious items of this scale were answered in such a way that they cancelled each other.

An interesting inverse relationship exists between the K scale of the MMPI and the PSEB scales. Stout (10) has concluded that K represents a plus-getting tendency in opposition to the traits in the other scales.

The higher mean scores found in this study for the PSEB when compared with the standardization group may indicate that the volunteer population supplying the original norms may reflect a selection factor. There are almost no differences between the normal male and female groupse xcept the expected significant Mf difference on the MMPI.

The MMPI profiles show a basic difference between normals and deviants to be an inverse F-K relationship. Normals tend to score low on F and high on the K scale while the deviant group scores in direct opposition to this [3, 8]. The profile with almost co-equal peaks produced by the deviant group on the clinical scales of the MMPI showed elevations at *D* and *Pt* and corresponds to the "basic anxiety profile."

Conclusions

1. Individuals showing psychological disturbances on the Minnesota Multiphasic Personality Inventory are also identified by the Psycho-Somatic Experience Blank.

2. The *Pt* scale of the MMPI has the highest linear relationship with both scales of the PSEB. Highest correlations between the PSEB psychological scale and the MMPI clinical scales, in order of relationship, are *Pt*, *Sc*, and *D*. The K scale of the MMPI also has a substantial negative correlation with this scale of the PSEB.

3. The clinical scales of the MMPI showing the highest relationship with the PSEB physiological scale in order of magnitude are *Pt*, *Hs*, and *Sc*. Again K has a definite negative relationship with this scale.

4. The other correlations between the scales of the MMPI and the PSEB show unsubstantial relationships.

5. Findings in the present study substantiate the hypothesis that K represents a factor opposed to "plus-getting."

6. Relationship between the *Hy* scale of the MMPI and the two scales of the PSEB may be of a higher order than indicated, but lowered by neutralizing factors in the way the group answered the subtle and obvious items of the *Hy* scale.

7. The total group mean scores for the normals on the MMPI seem to be representative of the male and female mean scores and reflect little in the way of sex differences.

8. There is a similarity of profile pattern among normal college groups as a whole, characterized by a higher *Hy* score than either *Hs* or *D* on the "neurotic triad."

9. The population herein studied is subject to such fluctuations as that produced by a heavy veteran loading, and a study should be conducted at a later date to measure for differences related to these factors.

BIBLIOGRAPHY

1. BROWN, H. S., Similarities and Differences in College Populations on the Multiphasic. *J. appl. Psychol.*, 1948, *32*, 541-549.
2. FRY, F. D., A Study of the Personality Traits of College Students, and of State Prison Inmates as Measured by the Minnesota Multiphasic Personality Inventory. *J. Psychol.*, 1949, *28*, 439-449.
3. GOUGH, H. G., Simulated Patterns on the MMPI. *J. abnorm. soc. Psychol.*, 1947, *42*, 215-225.
4. HATHAWAY, S. R., and MCKINLEY, J. C., *Manual for the Minnesota Multiphasic Personality Inventory*, N. Y., Psychol. Corp., 1943.
5. MCFARLAND, R. A., and SEITZ, C. P., A Psycho-somatic Inventory. *J. appl. Psychol.*, 1938, *22*, 327-339.
6. MCFARLAND, R. A., and SEITZ, C. P., *Manual for the P-S Experience Blank: Psycho-somatic Inventory*, N. Y., Psychol. Corp., 1947.
7. MCKINLEY, J. C., and HATHAWAY, S. R., A Multiphasic Personality Schedule; II. A differential study of hypochondriasis. *J. Psychol.*, 1940, *10*, 255-268.
8. MCKINLEY, J. C., HATHAWAY, S. R., and MEEHL, P. E., The Minnesota Multiphasic Personality Inventory; VI. The K scale. *J. consult. Psychol.*, 1948, *12*, 20-31.
9. MEEHL, P. E., and HATHAWAY, S. R., The K Factor as a Suppressor Variable in the MMPI. *J. appl. Psychol.*, 1946, *30*, 525-564.
10. STOUT, MARJORIE B., *An Analysis of the Structure of the MMPI.* Unpublished Master of Science thesis, The Pennsylvania State College, 1949.

THE CLINICIAN AND HIS PREDICTIONS

CHARLES A. DAILEY

St. Louis, Missouri

I.

Sarason [14] has recently raised the question: Why such poor predictions from clinical tests? His answer is that, in making predictions, we typically neglect certain critical variables, e.g. the nature of the instructions, the purpose of the testing as understood by the subject, etc. If the effects of these variables were known, the clinical psychologist could make predictions with a higher accuracy than at present. We shall inspect this proposed remedy more closely. There is no doubt that clinical testing is in an unsatisfactory condition. A reevaluation is plainly in order, and serious proposals for new lines of investigation always merit careful judgment.

Sarason's proposed remedy is "to identify the determinants of behavior, to state the principles by means of which they become related", and to put such theory to experimental test. However, conventional psychometric theory also postulates determinants. Intelligence test scores are determined by a postulated "intelligence" plus an "error factor"; Sarason calls into sharp question the adequacy of such a theory. He proposes to approach clinical testing with a methodology which is more congruent with current methodologies in the fields of learning and personality theory than psychometric theory. It is reasonable to hope that if his approach were systematically developed, a variety of puzzling phenomena in psychometric stituations would be resolved. The predictive power of formal test scores could conceivably be raised.

II.

It is precisely at this point that the limitations of such a proposal become manifest. Will his proposal raise the level of *all* predictions made in the clinical situation, or will its effects concern only the predictions of formal measurement? Any remedy for the state of affairs in the actual clinical situation must take account of this (probable) fact: In the clinic, it is not only the test score which predicts; it is also the clinician who makes predictions—and who fails.

Sarason's examples are illustrative of this point. He cites findings of the Assessment Program [11]—the generally low level of predictions of future clinical success.* These predictions were of *two* kinds: predictions by test *scores*; predictions by *clinicians* (using not only scores but raw data as the bases for predicting). Sarason's research could conceivably improve the predictions made by objective scores. It is more doubtful that it would necessarily improve predictions made by clinicians. His other example described a failure in appraising the personality of a child seen therapeutically. This failure did not involve test scores in the formal sense at all. It was a clinician who was 'taken in' by the effective defensive tactics of the child.

In our opinion, the remedy for poor clinical prediction is a better, a more understanding, clinician. We do not need to measure the determinants of test scores so much as we need to measure the determinants of the clinician's accuracy in understanding a patient. Examples of such determinants might be: (1) projective vs. objective data; (2) method of organizing a given body of data on the patient; (3) type of personality defenses and needs within the clinician. A particularly important question is whether knowledge of formal psychology really aids prediction. As challenging and disturbing as such a question may be, it is still relevant to ask it, as Luft [12] has done. While the author has data to suggest that it does aid prediction, Luft's data do not substantiate such an hypothesis.

The remedy we offer, in short, is nothing less than the comprehensive investigation of the clinician as the instrument of prediction. The practical goal of this investigation (while it also has innate theoretical interest) is to increase the clinician's ability to predict the patient's behavior—with or without tests. This calls for an investigation of the determinants of understanding. Unless it proves too embarrassing, the clinician should enter the academic laboratory as subject.

III

We propose that to phrase the question about clinical prediction in this manner is to meet the needs of modern clinical psychology as they stand. Some personality theorists have gone so far as to urge that it is necessary to develop a theory of how the clinician *and* patient interact [16]. This is essential, they claim, in order to understand the more important phenomena of clinical and therapeutic situations. Our proposal is considerably more modest. We would concentrate on the predictive accuracy of the clinician as the central variable. The conditions under which such understanding fluctuates would be the "independent variables". The proposal is not new. It does no more than take explicit recognition of a growing tendency toward this kind of research. The research studies the clinician as predictor rather than the test score. The author advances the opinion that such work, vastly enlarged and systematized, offers hope for bettering the unsatisfactory state of clinical prediction.

It might be said to include several broad divisions of investigation. First, the collection and presentation of clinical observations upon the problem (the clinician observes the clinician). Second, the formulation of hypotheses from such observations and their systematic disproof. Third, the development of psychologically and mathematically adequate measures. And fourth, the mature construction of systematic theory.

In the first division, the work of Theodor Reik [13] is widely known. From experience in the role of psychoanalyst, he has formulated the kind of suggestive, informal conjecture which is the necessary first step in any scientific theory. In the second division, the experimental work of Asch [3], Cartwright and French [4], Dailey [5], Dymond [6, 7, 8], Hanks [9], Isaacs, Fielder, and Fiske [10], Luft [12], Valentine [18], and Weingarten [19] are relevant. Most of these involve the method Allport described as "controlled cross prediction" [1] and also described by Steinmetz [15].

*In the present author's opinion, over-all evaluations of the success of the Assessment are grossly premature, inasmuch as its chief scientific value is as a genetic study. It was begun only four years ago.

The author knows of no substantial work on the methodological problem of measuring understanding. In general, there is a trend away from the use of reliability of judgments as an index of validity. There is also a trend away from the use of trait ratings (with the notable exception of the Assessment [11]) in this area of research, and a trend toward the use of predictions of known behavior as the criterion of understanding. The fourth area of contributions, that of the formulation of comprehensive theory, is largely devoid for obvious reasons. Allport's valiant attempt [2] to offer broad generalizations regarding the nature of understanding might be said to be premature. Nevertheless, such clearly stated theory is immensely valuable as a means of crystallizing research issues.

IV

In the course of the day's work, the average clinical psychologist makes many predictions—explicitly and implicitly. "I shall reassure Mr. Jones at this point in our interview" . . . "This child would benefit from play therapy today" . . . "This patient is going 'downhill' and should be put on closed ward." It is likely that a variety of predictions is made which far exceeds the variety of objective test scores and their known correlates. This seems true to the author in spite of the apparent attempt of psychologists to make a measuring stick for every possible kind of behavior. Whether we approve or not, the clinician functions as an instrument of prediction. We may exhort him to base all his predictions on objective measures if we wish. Or we may accept the necessity for non-objective clinical method and strengthen its validity by examining it under all its conditions of operation. We propose the latter.

We should carefully distinguish between objective measurement as a tool of scientific research and as a tool of clinical investigation. There is no question that it is essential in advanced scientific research. In fact, the research proposed in this paper assumes that understanding may be objectively measured. On the other hand, clinical investigation seeks only to predict; the means of prediction are irrelevant. All that is needed is a high level of prediction and, in that sense, clinical investigation is an "engineering" problem; any evidence is useful.

An outstanding example of confusion between the two aims of scientific and clinical investigation is the Rorschach research. A quantitative horror, the Rorschach has been used in objective investigation as well as clinically. This has apparently been due to the supposition that if the Rorschach is useful clinically, then it must contain behavioral segments susceptible to quantification. This supposition is justified only if we assume that the clinician functions like a scoring key. It is, of course, a plausible hypothesis that if the clinician does not predict from objective scores to future events, then he must be predicting from subjective scores. This view was offered by Thorndike [17] as theory thirty years ago; but to this author's knowledge, it has never been verified.

V

We are indebted to Sarason for his public discontent with the state of clinical prediction. We take issue, however, with his diagnosis of the difficulty and with his remedy. An extension of our scientific knowledge of the determinants of test behavior is commendable on theoretical grounds. But it is uncertain what influence this extension in formal knowledge will have upon the clinician's ability to understand the particular patient, although it will no doubt improve the accuracy of objective predictions.

Formal knowledge of such determinants *may* improve the clinician's understanding of the patient-in-particular, but this is not axiomatic. Fortunately this whole issue is subject to empirical test. To know what will increase the accuracy of the clinician's predictions means to know what the determinants of accuracy are. There is a growing body of information about how the clinician develops an impression of a

person and translates this impression into predictions of that person's behavior. This information is available on several levels of exactness. We have advanced in this paper the working hypothesis that the level of clinical prediction may be raised by scientific inquiry into the *conditions* under which high predictability is achieved. The focus of such research is upon the clinician.

Bibliography

1. Allport, G. W. The use of personal documents in psychological science. *Soc. Sci. Res. Coun. Bull.*, 1942, No. 49.
2. Allport, G. W. *Personality.* Henry Holt & Co.; New York, 1937.
3. Asch, S. E. Forming impressions of personality. *J. abn. soc. Psychol.*, 1946, 41, 258-290.
4. Cartwright, D., and French, J. Jr. The reliability of life history studies. *Character & Pers.*, 1939, 2, 110-119.
5. Dailey, C. A. *Some factors influencing the accuracy of understanding personality.* Unpublished doctoral dissertation, University of Michigan, 1950.
6. Dymond, Rosalind F. A scale for the measurement of empathic ability. *J. consult. Psychol.*, 1949, 13, 127-133.
7. Dymond, Rosalind F. A preliminary investigation of the relation of insight and empathy. *J. consult. Psychol.*, 1948, 12, 228-233.
8. Dymond, Rosalind F. Personality and empathy. *J. consult. Psychol.*, 1950, 14, 343-350.
9. Hanks, L. C. Prediction from case material. *Arch. Psychol.*, N. Y. 1936, No. 207.
10. Isaacs, K. S., Fiedler, F. E., and Fiske, D. W. *Some factors involved in the understanding of patients by clinicians.* Paper read at Midwestern Psychological Association meeting, 1950.
11. Kelly, E. L. and Fiske, D. W. *The selection of clinical psychologists, a progress report.* University of Michigan, Dec., 1948.
12. Luft, J. Implicit hypotheses and clinical predictions. *J. abn. soc. Psychol.*, 1950, 45, 756-759.
13. Reik, Theodor. *Listening with the third ear.* Farrar, Straus, & Co., New York, 1948.
14. Sarason, S. S. The test situation and the problem of prediction. *J. clin. Psychol.*, 1950, 6, 387-392.
15. Steinmetz, H. C. Measuring psychological understanding. *J. clin. Psychol.*, 1945, 4, 331-335.
16. Sullivan, H. S. *Conceptions of modern psychiatry.* William Alanson White Psychiatric Foundation, 1947. Washington, D. C.
17. Thorndike, E. L. Fundamental theorems in judging men. *J. appl. Psychol.*, 1918, 2, 67-76.
18. Valentine, C. W. The relative reliability of men and women in intuitive judgments of character. *Brit. J. Psychol.*, 1929, 19, 213-288.
19. Weingarten, E. M. A study of selective perception in clinical judgment. *J. Personality*, 1949, 17, 369-406.

DIRECTIONS FOR ADMINISTRATION OF THE ACHROMATIC-CHROMATIC H - T - P

JOHN N. BUCK

Lynchburg State Colony

While the H-T-P manual (which later appeared as Clinical Psychology Monograph No. 5) was in process of publication, we began to administer from time to time a chromatic H-T-P following the conventional achromatic H-T-P and the subsequent P-D-I. It soon became obvious that this extension of the H-T-P was most productively rewarding and it became routine procedure. Data accumulated over the past two years afford rather convincing evidence that we were fully justified in doing this, even if we were to consider only the obvious advantage of having an opportunity to compare drawings made at two different levels of frustration.

In order that clinicians may be acquainted with the several changes in administrative procedure which have taken place since the publication of the manual, the following brief description is offered:

STEP I:

Present the subject with the form sheet folded so that page 2 only is visible and the printed word "House" is at the top of the page. Ask the subject to take a pencil and draw as good a picture of a house as he can. Inform him that he may draw any kind of house he likes, he may erase as much as he wishes, he may take as long as he chooses, but he should draw as good a house as he can. Tell him to let you know when he is through.

If the subject protests that he is not an artist, assure him that the H-T-P is not a test of artistic ability at all. If he asks permission to use, or attempts to use, any mechanical aid, tell him that his drawing must be free-hand. Then, in turn, ask the subject to draw as good a Tree and as good a Person as he can, presenting the appropriate page of the drawing form to him in each instance, taking care to see that the printed word is at the page's top from the subject's point of view.

After the subject has been presented with the form sheet and has been told what to do, as examiner start your stop-watch, and using the first page of the scoring folder, record for the drawing of the House: (1) the amount of time that elapses between the time when you completed giving instructions to the subject and the time when the subject first begins to draw (we call this *initial latency*); (2) by name the details of the House, numbering each in the order in which it is drawn and taking care to be as minute and specific as possible; (3) any pause occurring once the subject has begun to draw, taking care to relate that pause to the detail sequence (these are called *intra-item pauses*); (4) all spontaneous comments made by the subject, relating them to the detail sequence; (5) any emotion exhibited by the subject, orienting that to the detail sequence, too; and (6) the total time consumed by the subject for his drawing. Repeat such timing and recording for the drawing of the Tree and the Person.

STEP II:

After the subject has completed his achromatic drawings, tell him that you will wish now to ask him a few questions concerning what he has just drawn so that you may be absolutely certain that you understand just what he was intending to portray in his drawings. The questioning should always be conducted as informally as possible, and any suggestion that you are criticizing the drawings should be studiously avoided. Using the Post-Drawing-Interrogation folder (1950 edition) ask the questions contained therein which are purposely spiraled to help prevent the establishment of an answer set. Then, since the P-D-I is intended to help the examiner ascertain what specific meanings and connotations the drawings have for the

subject, as well as to make sure that the examiner fully understands the subject's intent so that later quantitative and qualitative appraisal may be more accurate, ask such additional questions as seem to be indicated. In certain instances, it is advisable to ask the subject to indicate with an appropriate pencil-dot the location of the hip joint of the drawn Person.

To conclude Step II, request the subject to draw a sun and a line representing the ground in each of the pictures in which he has not already entered them spontaneously, and note which were induced.

STEP III:

Ask the subject to identify by name the color of the eight crayons before him (his inability to do this would make it imperative that the examiner test for color blindness),* then present him with a fresh form sheet with page 2 only visible, with the word "House" at the top of the page. Ask the subject to draw as good a House as he can, using any or all of the eight crayons**. Be sure that no pencil is available in this step and that no eraser is at hand, for the subject is not permitted to erase in Step III. Do not ask the subject to draw *another* House, since the subject may construe that as demanding that he draw a chromatic house which is unlike his achromatic house (the subject is always to be allowed complete freedom of choice in this respect).

After instructing the subject, the examiner will turn to page 4 of the scoring folder (1950 revision) and record, as in Step I, initial latency, time consumed, detail sequence (taking care to note not only the color used for each detail, but the color of any crayon or crayons tentatively selected but not actually used), emotional manifestations, and so on. Instructions for the drawing of the chromatic Tree and Person are identical with those in Step I, except that the subject is told that in this instance he may use any or all of the crayons at his command.

STEP IV:

After the subject has completed his crayon drawings, the examiner should ask whatever questions are needed (1) to make clear the subject's intent; (2) to identify and account for, as far as possible, conceptual differences between the achromatic and chromatic productions; (3) to make possible accurate quantitative and qualitative analysis. As a rule, we employ the following abbreviated P-D-I routinely, expanding it if time permits and the subject verbalizes freely.

PERSON

1. Is that a man or a woman (or boy or girl)?
2. How old is he?
3. Who is he?
4. What is he doing? (and where is he doing it?)
5. What is the weather like in this picture?
6. What does that person need most?
7. What kind of clothing does this person have on?
8. Show me where the person's hip joint is.

TREE

1. What kind of tree is that?
2. How old is that tree?
3. Is that tree alive? (if not, what part is dead and what caused it to die, *and* will the tree come to life again, and when?)
4. Which does that tree look more like to you: a man or a woman?
5. What is there about it that gives you that impression?
6. What is the weather like in this picture?
7. Show me the direction in which the wind is blowing.

---

*Attempts to interpret the use of color *per se* should always be made with the utmost caution, and should be restricted to subjects who are not color-blind to any marked degree.

**_Crayonex_ of the American Crayon Company, Sandusky, Ohio .

HOUSE

1. How many stories does that house have?
2. Whose house is that?
3. What room would you take for your own? Why?
4. Whom would you like to have live in that house with you? Why?
5. What is the weather like in this picture?

Next try to get the subject to identify any irrelevant details and to account for (1) the absence of "essential" details; (2) the presence of unusual or bizarre details; (3) unusual proportional, spatial, or positional relationships.

Have the subject draw a sun and a line representing the ground in each drawing in which he has not previously done so spontaneously, and note which were induced. Conclude by writing a brief description of the actual weather pertaining at the time of examination (so that the possible significance of "pressure" questions can be better evaluated) and by noting whether the subject is dextral or sinistral (so that profile presentations can be more accurately interpreted).

A TABULATION METHOD FOR ANALYZING COMBINATIONS OF RORSCHACH SCORES

KENNETH R. HAMMOND

University of Colorado

One persistent difficulty in clinical research involves the statistical problem of handling several variables simultaneously. Research with the Rorschach test is outstanding in this respect, and indeed, the problem can be generalized to most clinical work involving several test scores. Cronbach [1] has presented an attractive statistical and graphical technique for handling up to three variables simultaneously. An alternative procedure is outlined below.[1]

Condition 1. Describing a group with reference to their performance on (say) three variables (or tests).

Divide the variables (or test scores) into (say) three categories: High, Medium, Low. Set up a tabulation sheet as in Table 1. Note that the three variables are at the top of the tabulation sheet, each divided into the desired number of categories (in this case, three). Make an entry in the appropriate column for each individual who bears the appropriate characteristic. For example, suppose one has 3 variables with scores running from 0 - 75. Assume 0 - 25 to be the Low category, 26 - 50 the Medium category, and 51 - 75 the High category for all three variables. Thus an individual with the following scores, A = 20, B = 21, C = 22, would be placed in the first column on the left in Table 1, i.e., low on A, low on B, and low on C. In Table 1 there are entered 20 such individuals. Suppose an individual to obtain the following scores, A = 42, B = 15, C = 10. This individual would be medium on A, low on B, low on C. In the tenth column from the left there are entered 19 such individuals.

From inspection, then, one might note the following: (1) where the largest number of cases falls; for example, whether in a given case the pattern high on A, low on

[1] No claims for originality are made by the author. The author was first made acquainted with this technique by James Sakoda and George Kuznets who used the technique herein described in connection with sociological data, although not exactly in the manner presented here.

B, medium on C includes the largest number of cases; (2) what the trend is, e.g., whether, as in the example illustrated, there is a particular trend on Variable A from low to high, the same trend on Variable B, but the opposite trend on Variable C. With reference to Variable C one notes that for all values of A and B, the number of cases increases as C varies from low to high. Thus, one is able to make statements about the effect of Variable C, with Variables A and B controlled.

Consider A, B and C to be Rorschach test variables: let Variable A be number of responses (R), Variable B number of movement responses (M), and Variable C be sum C (ΣC). Referring to the illustration, one might then note how M and ΣC vary with respect to those cases low on R (say, 1-15 responses). In the hypothetical case, in the low R category it is apparent that an increase in M is related to *decrease* in the number of cases, while an increase in ΣC is related to an *increase* in the number of cases. The above illustration is intended to present the general method and deserves only slight consideration because it does not indicate how one might discriminate between groups—the latter, of course, being one of the important problems in connection with the Rorschach and other tests.

Before discussing the discrimination problem, it should be noted that the number of cases available is the principal limitation on the number of variables which can be handled. The writer has participated in a research project[2, 3] which included 109 cases and found it feasible, though not completely satisfactory, to use five variables. Some variables were dichotomized into "presence" and "absence" categories, as for example, presence or absence of "rejection" of certain Rorschach cards.

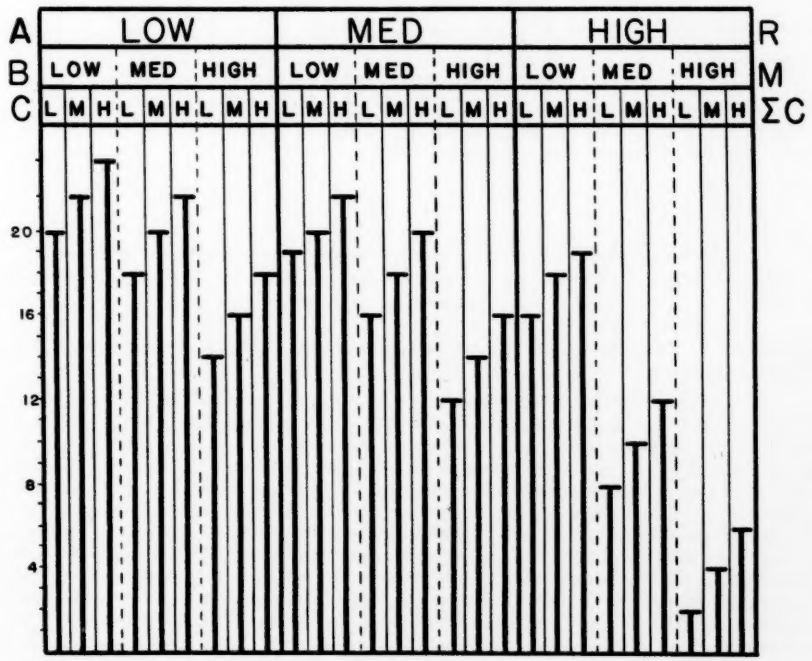


TABLE 1. ILLUSTRATING THE METHOD FOR GRAPHIC COMPARISON OF SEVERAL VARIABLES SIMULTANEOUSLY.

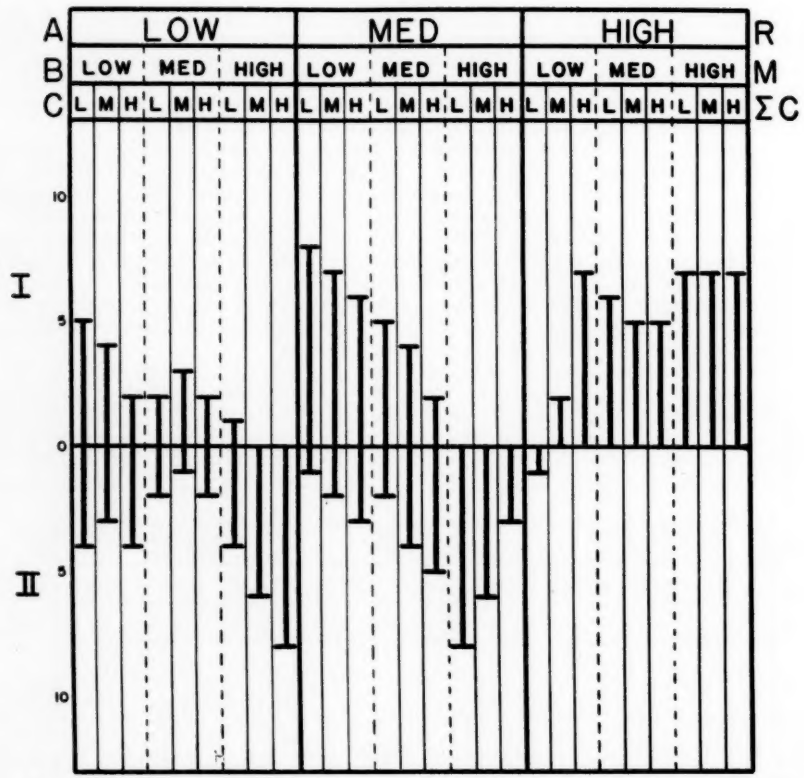


TABLE 2. SHOWING A METHOD FOR COMPARING SCORES OF TWO GROUPS (I AND II) ON SEVERAL VARIABLES SIMULTANEOUSLY.

Condition 2. Discriminating between two groups with reference to their performance on three variables.

In this problem the table is set up in the same manner as Table 1, with the exception that a line is drawn horizontally across the table as in Table 2. *Above* the line are entered those individuals in Group I (patients, say) who bear the appropriate characteristic, and *below* the line are entered those individuals in Group II (normals, say) who bear the appropriate characteristic. It is thus possible to observe which combination of variables appears in one group and not the other. For example, in Table 2, (again let A = R, B = M, C = ΣC) consider first R; there seems to be about as many cases above as below the line except in the category High R. Apparently, few members in Group II bear the characteristic "high number of responses." Those individuals in Group II who do fall in the High R category can be described with reference to M and ΣC. Consider M; note that for *low* and *medium* values of R, low number of M responses appear with members of Group I, but that a *high* number of M responses appear with members of Group II. Within the group which shows a high number of responses, variability in M is not clearly a differentiating variable, except to the extent that there are certainly more individuals in the High M category than in the Low M category. Consider ΣC: in the low R group, variability in ΣC

from low to high seems to increase the number of cases in Group II with respect to those individuals with High M. In contrast, note that in the high response category, variability in ΣC seems important only in regard to those individuals with a low number of M responses.

The above statements serve only to indicate the kind of information one might obtain from such a table. *In general*, the following information might be obtained: (1) what is the effect of variation in each variable considered separately with the others controlled, e.g., what is the effect of varying M with R and ΣC controlled? (2) what is the effect of all three variables considered as trends? (3) what is the effect of certain combinations of scores? The foregoing illustrates the way the table might be used as a way of getting "hunches" about the data, i.e., as a "hypothesis-producer." Various combinations of variables will be turned up which can then be examined in the light of theory.

While examination in itself is important in that few techniques exist for examining the effects of several (more than three) variables simultaneously, a test of significance combined with the tabular method would be extremely valuable. Both Cronbach's method and the method described here provide information about the frequency of occurrence of patterns with respect to the individuals (not average scores for groups) of two different groups. Significance tests may then be applied to the occurrence or non-occurence of these patterns in the criterion groups.

Summary

A tabulation method of handling several Rorschach variables simultaneously is presented. This method has been tried in practice and found to be economical and useful.

References

1. Cronbach, L. J. Pattern Tabulation: A Statistical Method for Analysis of Limited Patterns of Scores, with Particular Reference to the Rorschach Test. *Educ. Psychol. Meas.*, 1949, *9*, 149-171.
2. Rogers, L., Knauss, J. and Hammond, K. R. "Prediction of Response to Therapy. "I. Desire for treatment." (In progress).
3. Rogers, L., Knauss, J. and Hammond, K. R. "II. Success in treatment." (In progress).

USE OF THE BLACKY PICTURES WITH A CHILD WHOSE OEDIPAL DESIRES ARE CLOSE TO CONSCIOUSNESS

H. MICHAL-SMITH, EMANUEL HAMMER AND HERMAN SPITZ

New Jersey State Diagnostic Center
Menlo Park, New Jersey

Introduction

Since Freud wrote "The Phobia of a Five Year Old", clinicians have become increasingly aware that children may identify more easily with animals than with humans. This awareness has cumulated in two important clinical instruments, both appearing in 1949. Bellak [1] developed his C.A.T., and users of this technique have since assembled empirical evidence that animals are preferred identification figures for children. His projective test is structured to tap the dynamics of inter-personal relationships, drive constellations, and the defenses employed against them. Dynamic factors of the child's group, home, or school reactions are explored. This differs from Blum's "Blacky Pictures" [3] in that the latter are more specifically conceived to fit psychoanalytic concepts. The cartoons employed are structured to explore oral eroticism and sadism, fixations, Oedipal intensity, masturbation guilt, castration anxiety, positive ego-ideal, love-objects, etc. The Blacky Pictures were originally intended for use with adults. Animals were used to circumvent resistances and facilitate freedom of projections in situations where human figures might provoke inhibitions "as a result of being 'too close to home'."

Our experience, however, has shown the Blacky Cartoons to be even more productive with children than with adults. Children have been found to throw themselves into this testing situation with much alacrity and enjoyment. As a result, an abundance of dynamic material is often elicited. The cumulative experience of many clinicians, working in varied installations, is necessary to determine the ultimate worth of a projective technique. It cannot be determined in a tight little study as can pencil and paper tests of the more objective variety. As an illustration of the efficacy of the Blacky Test with children, and as a stimulus to other clinicians to use and assemble data with and on this tool, the present case is presented.

Case Study

Background. The patient is a 9 year old negro boy, referred to the Diagnostic Center for stealing toys. There is also a history of spying on his mother and step-father at night, masturbation, homosexual and heterosexual episodes. He shares a bed with two brothers, one 7, and one 13. There is an older sister of 18. The client is a rejected child, and has only once in his life received a present from his mother. He does not know when his birthday is, due to his mother's habitual ignoring of the event. As punishment, he is often deprived of food, and when he is enuretic he is put in cold water to be taught a lesson. He feels rejected by his teacher as well as his mother, and is beaten up by his older brother while his mother is at work. Hence, in his self-concept, he unconsciously thinks of himself as a "football dummy" whose function, we may hazard a guess, is only to absorb punishment.

Blacky Projections. The most productive cards were II (Oral Sadism), XI (Love Object) and IV (Oedipus Intensity), although all the cards produced fertile responses. "Mama's collar", which Blacky is biting in Card II, turned up in Card III, IV, V, IX, X and XI. This is a marked persistence of an object drawn into a story and is indicative of the importance that the patient attached to the retaliatory wishes against his mother and his guilt over this. Some excerpts of his carry-over responses are:

---

*H. Michal-Smith, Chief Psychologist, New Jersy State Diagnostic Center; Emanuel Hammer and Herman Spitz, Psychological Interns.

Card III (Anal Sadism). "He's probably covering it and when he's through he's going back to tearing the leash, and his mother, she finds out, she might keep him in the house for 2 days without food." (Strong super-ego development and equating of food and love show up throughout.)

Card IV (Oedipal Intensity). "Papa probably saying that "He tore the leash up, I'll have to get you (Mamma) another." Blacky getting mad because he said that, because she's getting a new one and he wants a new one, but mommy and poppy won't get him one. Because he tore up the old one; when mommy first met poppa he gave her a brand new one and it got old so Blacky tore it up." (Castration feelings, oedipal jealousies, desires for a more adequate penis (as adequate as father's), and vindictive reactions to concept of mother's incorporation of father's penis all gleam through the emotional mosaic of this story.)

Card V (Masturbatory Guilt). "Blacky probably forgot about his sex and didn't even know about it, and while he was chewing the leash and watching mama and papa all night, and listening to papa and mama saying they would give the brother all the milk, and he forgot all about the sex, and then he remembered it. Then he started to clean his sex (penis) . . . (later) . . . Then he'll take a leash from his father, and brother, and all 3 will punish him. Then he'll have to tear up his leash, and he'll have to punish himself." (Feelings of rejection and envy of younger sibling, primal scene evidence, masturbation as a reaction to oedipal jealousies, guilt rising to castration proportions, and oral aggression and fixation are to be found in this rich story. Note shifting symbolism of leash in all these projections.)

Card IX. "He stays out scared. His brother Tippy don't think he deserves the leash."

Card X. "The first day he ripped up leash Mother didn't want him there . . . If he would come back he wouldn't get nothing but punishment."

Card XI. "Probably dreaming he's making love to Mama, and that he bought her a new leash and Papa is home sitting in his coop, just as Blacky was." (Phantasy consumation of oedipal wishes are made possible by concomitant phantasy acquisitions of a new and adequate phallus. As in Freud's case of little Hans, he is impatiently anxious for his penis to grow at a more rapid rate.)

The last response is indicative of the close to conscious Oedipal feelings that the patient expresses, and which pervade the record even as the "leash " does. It also reveals the ambivalence the patient feels. The collar is torn up because his love is not returned, but with amends (a new collar) he takes the place of his father. Now for the specific Oedipal responses.

Card IV (Oedipal Intensity). (*laughs*). "Just because Mama and Papa is making love Blacky is getting mad and getting angry. Probably getting mad because he has no one to love."

Card V (Masturbatory Guilt). "And while he was chewing the leash and watching Mama and Papa all night . . . " (Chewing the leash is equated with masturbation while entertaining sex phantasies involving mother.)

Card VII (Positive Identification). *Inquiry No. 4:* (The question is: Whom would Blacky rather pattern himself after (be like), Mama or Papa or Tippy?)
Answer: "Papa, cause Papa can make love to Mama but Tippy can't."

Card X. *Inquiry No. 1.* (The question is: Whom does the figure remind Blacky of?)
Answer: "Him running away . . . Think it's better to run away than stay home and get punishment. Thinking of home and Papa making love to his mother. Thinking why couldn't he be making love to his mother."

Card XI. "Probably dreaming he's making love to mama."

There were many other provocative responses, especially concerning the sibling rivalry. There were numerous slips, at which time Blacky was called Tippy. This occurred on Card VI (Castration Anxiety) where Blacky's tail was cut off—not Tippy's. It is interesting to note here that Blacky also threw the knife. Another time he confused the two names on Card IX (Guilt Feeling): "Tippy's probably getting scared cause his Mom isn't paying attention to him . . " Again on Card VIII (Sibling Rivalry): "Brother probably did something good and they're petting Blacky." In the first example, castration fears forced their way out. In the latter two examples, the patient traded positions in phantasy with his favored brother.

It is interesting to note the part that feelings of race inferiority can play in these projections. In actuality, the favored sibling was lighter skinned:

Card I. (Oral Eroticism). "They probably don't like him. Because all the others are black and white and he was black."

Interpretation. In conclusion then, the Blacky Pictures reveal strong ambivalent feelings towards the mother which covered an underlying love for her and concomitant hatred of the father. When the love was not completely returned, and was shared with a sibling, the negative side of the ambivalence arose. Hostile impulses are then directed toward the mother (and her collar) and pervade the patient's record.

Emotionally, he lives in a world of rich imagination and inner life. Loss of parental affection and marked feelings of rejection have caused anxiety, insecurity, and feelings of inadequacy. He has utilized his creative imagination to attempt to regulate and stabilize his relationships with the most important persons in his environment. The bulk of this he achieves in his phantasy life. Rejection has caused him to become overly introverted for his age and to withdraw into an inner life to obtain the satisfaction he feels inadequate about obtaining from the environment.

The fear of the consequence of acting out his emotional needs center around the psychosexual area. This area is a plethora of conflicts, including phallic preoccupation, masturbatory guilt, feminine identification, concept of the penis and homosexuality as something fearful. He views his sexual desires as that of a "beast" and this is probably due to severe punishment imposed when his mother discovered him having intercourse with a girl at the age of 5, and again caught him in the act of mutual masturbation with his younger brother at the age of 7. While his sexual guilt is reinforced by these acts and consequential chastisements, the original guilt springs from the fact that he is enmeshed in intense oedipal wishes which bubble dangerously close to the surface of his personality. There appears to have been an early traumatic situation which involved his mother and his step-father, and the subject is angry because (as he projects on one card) "Papa can make love to Mama and the child is thinking why can't he be making love to his mother." This anger figures as a factor in his stealing and misbehavior as an effort to punish his mother, and at the same time to afford him attention. The Blacky projections indicate that he resorts to masturbation because of his mother's love for another man, a father figure, and because of the younger brother's more secure spot in his mother's affection. His stealing represents the acquiring of the toys, and symbolically of the love, his mother has always withheld from him. The equating of toys, food, and love, and their deprivation, leave him with an insatiable desire for all three.

A word of caution: The Blacky Cards are especially constructed to elicit certain responses. Only the strength of responses, and their persistence, should govern the importance placed upon them. Experience in their use is the best guarantee.

Summary

Our clinical experience indicates that the Blacky Pictures are very effective in stimulating projective productions in children. A case is cited in which oedipal desires were quickly revealed in a 9 year old Negro boy who was referred for stealing toys.

Bibliography

1. Bellak, L. The Concept of Projective Psychology. *Projective Psychology* (Eds. L. E. Abt & L. Bellak). New York: Knopf, 1950.
2. Bellak, L. On the Concept of Projection. *Psychiatry*, 1944, 7, 353-370.
3. Blum, G. S. *The Blacky Pictures:* Manual of Instructions. New York: Psychological Corporation, 1950.
4. Blum, G. S. A study of the Psychoanalytic Theory of Psychosexual Development. *Genet. Psychol. Monogr.*, 1949, 39, 3-99.
5. Fenichel, O. *The Psychoanalytic Theory of Neurosis.* New York: Norton, 1945.
6. Freud, Sigmund. *Collected Papers.* London: Hogarth Press And The Institute Of Psycho-Analysis, 1948, 149-289.

THE APPLICATION OF HYPNOSIS TO NONDIRECTIVE PSYCHOTHERAPY*

MILTON V. KLINE

Westchester County Department of Health, Division of Mental Hygiene

Introduction

The use of hypnosis in psychotherapy is frequently reported as an aspect of analytically oriented therapy such as hypnoanalysis [13] or in a variety of directive hypnotherapy procedures [12]. In addition, recent work in the area of hypnodiagnostic testing and vocational analysis has emphasized the value of hypnotic states for personality investigation and psychotherapy [5, 6, 7, 8, 9, 10]. In both diagnostic and therapeutic applications of hypnosis there is a growing recognition of the fact that a light hypnotic trance can be meaningfully incorporated into psychological and psychiatric work [1, 2]. The ease with which such a state can be obtained in a large majority of patients significantly extends the applicability and value of hypnosis in clinical areas.

Hypnosis is usually considered a directive, clinician-oriented procedure and not readily associated with nondirective techniques. This consideration is largely an assumption based upon a limited evaluation of the various techniques and applications of hypnosis. The induction of hypnosis can be considered a patient-centered procedure when the approach to induction emphasizes the role of the patient as an active and productive element in the achieving of the trance state. Within this orientation, the use of indirect induction methods [13] completes the framework of a nondirective approach to the therapy relationship.

Recent work by Guze [3] has pointed to the fact that the induction of hypnosis can be utilized by itself as a diagnostic technique which leads to significant insights into personality drives and organization. Schneck [11] in reporting his use of hypnotic techniques in psychotherapy notes the spontaneous arousal of motor phenomena which relate to the patient's personality structure and emotional conflicts. The products of spontaneous reactions to and the handling of the hypnotic induction procedure can be incorporated into both the diagnostic and therapeutic processes involved in clinical psychological and psychiatric work.

In psychotherapy, hypnosis and hypnotic techniques may be used systematically throughout the course of treatment or may be used at precise times for specific purposes. The use of hypnosis systematically as the major operational framework is to be found in hypnotherapy itself, directive or nondirective, while the use of specialized hypnotic techniques is to be found more generally in psychotherapy where hypnosis is used for specific exploration, evaluation and productivity. The choice of either approach would depend largely on the therapist and the patient's presenting problems.

The purpose of this paper is to present several illustrations from clinical practice which describe the use of specialized hypnotic techniques within a framework of psychotherapy or psychodiagnosis characterized by a nondirective orientation and not essentially hypnotic in nature.

In Eliciting An Anamnesis

The use of a nondirective technique in taking a case history has many diagnostic advantages not found in more directive procedures. For one, it places greater emphasis upon the patient's spontaneity, and the manner in which he organizes and presents his spontaneous recollections brings significant psychodynamic elements into focus. This initial use of nondirective technique also sets from the first clinical con-

*Read before the Society for Clinical and Experimental Hypnosis, December 1950.

tact the orientation to be utilized in psychotherapy. In presenting the patient with a less delimited frame of reference within which he can locate his social, vocational and personal chronology, the possibility of self-initiating associations is enhanced. Thus, it is not an infrequent experience of a nondirective therapist to find that his first few clinical contacts with the patient bring forth through nondirective methods a more meaningful elaboration and portrayal of life history than was obtained during a preceding directive anamnesis.

Frequently patients have difficulty in organizing their memories or thoughts and their mechanisms of resistance and defense often result in guarded or severely circumscribed productions of life history with the suppression of meaningful associations. There may, in some cases, be difficulty in stimulating self-productivity because of anxiety, fear, and uncertainty of the therapeutic role. This is not infrequent since transference relationships do not usually come into being rapidly enough to be of productive value during the first few sessions. It is this problem among others that has made the psychological testing procedure valuable since pertinent material can be elicited in that situation without a transference relationship.

Hypnosis per se appears to produce immediately a transference relationship of its own which, while it may undergo change during the course of therapy, will nevertheless establish a situation in either the diagnostic or therapeutic contacts which tends to stimulate unconscious productivity and tone down ego defenses [9, 10]. This phenomena achieved through the induction of hypnosis per se stimulates patient productivity through releases of imagery activity, increased recall, spontaneous associations of unconscious origin, and motor phenomena linked to resistances, fears, and impulse characteristics.

In this connection a thirty-year old veteran seeking psychotherapeutic help for what he termed his "marital and personality problems" will be described. The patient, a man of superior intelligence, had previously been seen by an intake worker for the purpose of taking a case history and clarifying the nature of the patient's problem. The material which was elicited in that essentially directive contact was meager, guarded and quite stereotyped, with little productivity of psychodynamic significance. In addition, the patient was described as being hostile to the procedure and while compliant in many respects was obviously negativistic in general.

The writer next saw the patient for the purpose of psychological examination prior to psychotherapy. In view of the material presented by the intake worker, it was decided to use a nondirective approach with this patient in taking again an anamnesis through the use of a specialized hypnotic technique.

When seen the patient was tense, anxious and generally impatient for "treatment". His anxiety as displayed in motor restlessness was pointed out to him with mention of the fact that in his present contact most of the activity produced would have to be forthcoming from him. His ability to regulate his own motor responses offered an easy access to the problem of self-initiated activity and control. A period of prolonged silence followed during which the patient relaxed considerably and his motor restlessness all but vanished. He smiled when he noticed this and said, "I see what you mean." At this point it was explained that further freedom in both psychological and physiological expression could be gained through increased relaxation. Following this, an indirect hand levitation technique [13] was used for the induction of hypnosis. A light trance state was obtained in about five minutes and the patient was informed that he could deepen this state of relaxation through self-activity and that a period of time to be decided upon by him was available for that purpose. He was told that when he was as deep as he wished to be in his state of relaxation he could tell the examiner so. In about three minutes the patient said that he was very relaxed and quite "deep". He appeared to be in a light-medium hypnosis and no attempt was made to deepen this state by the examiner. Instructions were then given for the production of scene visualizations [5, 8] which the patient was able to satisfactorily produce with his eyes closed. The patient was then informed that he would produce a series of scenes which would describe his life history.

The patient immediately began to describe scenes from his boyhood and for the next eighty-five minutes spoke without interruption. In very coherent chronology he described the events leading from his boyhood to his present visit to the clinic. Some of the scenes were happy ones and the patient's bodily reactions reflected this, others were traumatic and again bodily expressions were correlated. Many of the scenes came as a surprise to the patient and seemed to be quite spontaneous. For example, he described a scene from his experience in military service which is detailed below verbatim:

> "I am walking out of a lobby of a hotel in — Texas and that WAC I told you about is with me. I can see her quite clearly but I am confused because while I recall this scene, the head and face of the girl in the picture is not the one I was with. You see, it's this WAC's body, but it's the face of that girl I told you about in New York. I don't understand that."

Following this production the patient was informed that he could produce associated scenes which would "explain the reason for the different face on the WAC".

This stimulated a production regarding the face which the patient associated with another girl, whom he had fantasies of having watched in sexual intercourse with another man. Elaboration of this productivity through additional scenes revealed significant material relating to a homosexual conflict which the patient had up to this time guarded even in his projections. He was able to venture some interpretation which indicated a grasp of some insights into his inner conflicts. Other productions of psychodynamic value were obtained in different areas of the patient's thinking and in general the case history which was elicited in this one contact was exceptionally full and inclusive of material pertinent to the patient's underlying personality difficulties. In addition to the diagnostic value of this material much of it could readily and easily be incorporated into psychotherapeutic procedure.

Following the termination of hypnosis the patient appeared to have good rapport with the examiner and expressed the idea that he felt much better and that real progress was being made in attacking his difficulties. He indicated that he was anxious to continue and expressed strong positive motivation toward the entire clinic contact. This incident indicates the value of an hypnotic technique in a nondirective anamnestic procedure which aids in the production of significant material and sets the pattern for increased motivation based on an emergence transference relationship within a patient-centered orientation.

Induction of Hypnosis As A Projective Technique

A fourteen-year old boy was referred to a mental hygiene clinic because of academic failure and "unhappiness". Diagnostic psychiatric and psychological examinations revealed a youngster troubled with neurotic anxiety which resulted in difficulty in all inter-personal relationships and particularly familial ones. A compulsive, stubborn, and generally defensive personality structure was noted. Plans were made for a psychotherapeutic approach involving remedial work in educational areas as well as help in personality adjustment. Initial contacts with the patient had been quite unproductive and the patient indicated little understanding of the need to come to the clinic and consequently little motivation to become involved in treatment.

As this became more apparent, an approach emphasizing the nondirective elements already instituted with the addition of hypnosis was decided upon. A visual fixation technique was used for induction, which the patient quickly associated with hypnosis. Through verbal comment and spontaneous motor behavior it became apparent that the patient was resisting hypnosis most intensely though he insisted that he was being cooperative. After thirty minutes a light trance with eye closure was obtained.

The patient produced little material while in hypnosis and in general acted out a pattern of extremely negativistic and hostile behavior which was in keeping with

his own underlying emotional drives and personality difficulties. Having achieved an hypnotic induction in the face of strong resistance, the patient now expressed his negativistic and hostile feelings toward the therapist by refusing to open his eyes when hypnosis was terminated. The full details of this pattern of reaction will be discussed elsewhere and will not be presented here in detail [4]. The therapist tried for over one hour to have the patient open his eyes. It was clear that the patient was frightened over this predicament because he insisted that he could not open his eyes and he tried very hard indeed to do this, including attempts at prying them open with his fingers. As time went on the patient's anxiety increased and he now was most cooperative in complying with all techniques designed to permit eye opening based on psychodynamic interaction. This even included the use of a colleague in an attempt to transfer the hypnotic relationship but which like other attempts failed to produce eye opening.

In view of the late hour and the patient's mounting anxiety a psychological shock technique was instituted. The patient was asked to stand up and face the door of the office. As he did so, the examiner walked quietly up behind him and suddenly clapped his hands very loudly right behind the patient's ear. The patient's eyes flew open and remained so. Thus, the first hypnotic interview was concluded.

The next interview found the patient considerably more cooperative and less resistant. He suggested a discussion of the difficulty involved in the previous session which he was still somewhat anxious about. For three following sessions the patient carried the bulk of the conversation and controlled its direction. He began to recognize in his hypnotic behavior signs of negativism and hostility. He couldn't account for it since he readily expressed the idea that he had no reason to dislike the therapist. Gradually he gained insight into the fact that he had handled that situation in the same manner that he handled many similar inter-personal situations and that there was in fact a transfer of reaction. From this point on, the patient gained increased motivation for psychotherapy and in a period of time which covered eight months he was able to increase his educational achievement level as measured by the Stanford Achievement Test from a grade status of 7.8 to 11.0. In addition, he decided to go to summer school and did so with success enough to warrant the school's accelerating him a full year in academic standing. His social relationships and family adjustment improved significantly.

In this illustration, an hypnotic approach served as a means for focusing the patient's thinking and feelings on his own drives and attitudes. The projective behavior resulting from this technique produced spontaneous activity in the patient which led to a clarification of his need for psychological help. The therapeutic relationship was accelerated and stimulated patient-centered motivation which led to successful psychotherapeutic procedure.

The Use of Hypnosis in Stimulating Associative Projectivity

A twenty-three year old married woman sought help in connection with her inability to hold a job and particularly her difficulties in accepting supervision. Vocational counseling revealed interests and abilities consistent with vocational choice and with jobs actually held. Attempts to relate personality factors to the vocational problem met with compliant stereotypy and a guarded resistance on the part of the patient. After eight counseling sessions during which little clarification or insight had been arrived at an hypnotic approach was decided upon.

A scene visualization technique devised for the investigation of unconscious motives involved in the choice of an occupation and in difficulties of vocational adjustment was employed [5]. Among many scenes involving vocational activity the following was elicited:

Patient: "I am in the office and Mrs. ——, the Office Manager, is calling me. I don't want to go but I will have to. I hate her."

Examiner: "Why?"

Patient: "I don't know why."

Examiner: "If you will look at the scene more closely you will be able to find out why."

Patient: "Oh! I never would have guessed that."

Examiner: "Guessed what?"

Patient: "Why Mrs. —— has my mother's face."

This production led to other productions which revealed a longstanding conflict between the patient and her mother which had been suppressed rather rigidly. From this point on, the counseling approach emphasized the problem as it had unfolded and led to sufficient insight and readjustment to permit a good vocational adaptation as well as an improved level of interpersonal relationship within the family structure.

This case illustrates the utilization of an hypnotic technique in vocational problems which are symptomatic of emotional disturbance. Resistance toward psychotherapy is frequently characterized by the patient channeling his complaints into a "non-emotional" area and putting up barriers of defense against the possible inclusion of personality maladjustment as the fundamental problem. In this case early counseling sterility was replaced by psychological productivity which led to a direct involvement of the patient on a level of therapy consistent with her problem.

Summary

This paper has presented several illustrative instances in which hypnosis combined with nondirective technique has led to significant diagnostic and therapeutic productivity. As an operative procedure for eliciting an anamnesis, in eliciting projective elaboration of psychodynamic behavior, and for stimulating associative activity, the use of specialized hypnotic procedures within a patient-centered setting appears to foster clinician-patient transference, stimulate self-productivity, and increase motivation for treatment. The material elicited through such procedures tends to be pertinent psychologically and can readily be incorporated into the total approach to the patient. In this respect hypnosis appears to fit into a nondirective therapeutic framework with procedural consistency. It in itself can be considered a nondirective method and can be presented and utilized in that manner. In view of this fact and since a light hypnotic state is useful for much clinical work, further utilization and investigation would appear to be warranted.

References

1. Ambrose, G. Hypnosis; its value in child guidance. *Brit. J. med. Hyp.*, 1949, 1, 42-44.
2. Ambrose, G. Child guidance and the first stage of hypnosis. *Brit. J. med. Hyp.*, 1950, 2. 41-43.
3. Guze, H. Wish fulfillment as an aspect of hypnotic induction. (to be published).
4. Kline, M. V. Negativism and the apparent loss of voluntary function following the induction of hypnosis. (to be published).
5. Kline, M. V. and Schneck, J. M. An hypnotic experimental approach to the genesis of occupational interest and choice. I. Theoretical orientation and scene visualization. (in press) *Brit. J. med. Hyp.*
6. Kline, M. V. and Schneck, J. M. Hypnosis in relation to the word association test. (in press) *J. Gen. Psychol.*
7. Kline, M. V. An hypnotic experimental approach to the genesis of occupational interests and choice. II. The thematic apperception test: a case report. (in press) *J. Gen. Psychol.*
8. Schneck, J. M. and Kline, M. V. Hypnotic scene visualization and the word association test. (in press) *J. Gen. Psychol.*
9. Schneck, J. M. and Kline, M. V. Hypnodiagnosis and evaluation of therapy in psychiatry and clinical psychology: report of a case involving the H-T-P. (in press) *Brit. J. med. Hyp.*
10. Schneck, J. M. and Kline, M. V. The H-T-P and TAT in hypnodiagnostic studies. (in press)
11. Schneck, J. M. Psychosomatic reactions to the induction of hypnosis. *Dis. Nerv. System*, 1950, 11, 1-4.
12. Watkins, J. G. *Hypnotherapy of war neuroses.* New York: Ronald Press, 1949.
13. Wolberg, L. R. *Medical hypnosis.* Volume I. New York: Grune and Stratton, 1948.

A NOTE ON EXTENDED FINDINGS WITH THE MMPI IN PREDICTING RESPONSE TO ELECTRO-CONVULSIVE THERAPY

JOHN S. PEARSON AND WENDELL M. SWENSON

Minnesota State Division of Public Institutions

In an earlier paper [1] evidence was presented to indicate that three signs in the pretreatment MMPI profiles of schizophrenic patients enabled good prediction of the response to electric shock therapy. Recently an analysis of the pretreatment records of 10 male and 23 female psychotic patients at the St. Peter State Hospital was carried out with respect to the prognostic signs discovered in the earlier study. The 33 cases, all first admissions, were rated by ward nurses as to the degree of improvement following electric shock therapy. The following breakdown with respect to diagnosis and results of therapy was observed: Ten patients were rated as "unimproved". This group included five paranoid schizophrenics, four schizophrenics of the hebephrenic type, and one psychopathic personality with psychosis. Three patients were rated as "slightly improved" including two schizophrenics of the paranoid type and one hebephrenic. Seven patients rated as "moderately improved" included three paranoid schizophrenics; one psychopathic personality with psychosis; one involutional psychosis; one manic depressive, manic; and one manic depressive, mixed. Thirteen patients rated as "marked improvement" included two paranoid, three simple, and one hebephrenic schizophrenics; three involutional psychoses; two manic depressives, manic; and two manic depressives, depressed.

While the small cell frequencies made a legitimate statistical analysis of the 4 x 4 fold contingency table impossible, inspection of the ungrouped data revealed that cases for whom all three prognostic signs were positive generally were rated "unimproved" following treatment, those with one or two signs positive improved slightly or moderately and those for whom none of the three signs were positive generally were rated "marked improvement". One case for whom a poor response to treatment was predicted since all three prognostic signs were positive actually was rated as moderately improved. This one instance differs with the findings from the Veterans Hospital group where no one for whom all three prognostic signs were positive was rated as having made an improvement. Folding this distribution into a 2 x 2 fold table yielded a X^2 value of 9.91 significant at the .01 level and a contingency coefficient of .48.

A scoring key for the Multiphasic had previously been developed by item analysis of the two groups of patients utilized in the earlier study. When applied to the groups from which it was derived this scale gave a very remarkable separation, and we were hopeful that it might prove more valuable in predicting response to treatment than the signs reported. However, when this scale was applied to the St. Peter State Hospital groups the results were most disappointing—the difference between the means for the "unimproved" and "marked improvement" groups barely approached statistical significance and the dispersion of individual cases indicated the scale to be useless in individual prediction.

Summary

Further evidence has been found that the Minnesota Multiphasic Personality Inventory may be of use in predicting the response of psychotic patients to electric shock therapy. Item analysis leading to the development of a new and specific scoring key proved less fruitful than examination of scores on the nine original clinical scales.

Bibliography

1. Pearson, John S. Prediction of the Response of Schizophrenic Patients to Electro-Convulsive Therapy. *J. Clin. Psychol.*, 1950, 3, 285-287.

NORMS FOR SCATTER ANALYSIS ON THE WECHSLER INTELLIGENCE SCALES

BENJAMIN ALIMENA

Fordham University and Catholic Charities Guidance Clinic

Introduction

The variability hypothesis for differential diagnosis rests on the assumption that similar numerical scores on the eleven subtests of the WBIS have the same quantitative value. The scores must be comparable. If this need is not met, then variability may actually exist even though scores are identical, and even when scores differ 'real' variation may not exist. It seems evident that this assumption is crucial and must be fulfilled, otherwise there will be no basis to compare scores for variability in abilities. But the most recent reports emphasize that this comparability of scores has not been achieved. Wechsler [4] used only the data of the 20-34 year old group of the standardization population to transmute raw scores into a weighted score scale. The variation between means and standard deviations is more pronounced outside this group. Thus in order to arrive at an adequate evaluation of the diagnostic function of the scale and of the variability hypothesis as an aid to differential diagnosis, comparability of scores must first be established. The purpose of this paper is to show that there is need to restandardize the weighted scoring system of the Wechsler-Bellevue Scale, and to outline a solution.

Method

The following formula is submitted as a solution for the establishment of comparability of scores, resulting in norms to evaluate the deviation of performance on the Wechsler sub-tests.

I. All the possible weighted scores for each of the sub-tests, for each of the age groups are transmuted into a standard score scale. Means and standard deviations are drawn from Wechsler's standardization data (*4*, table 39, pg. 222) so that the norms established can be considered reflective of the standardization population, (*Table 1, col. 3*).

II. The standard scores are now transformed into an equivalent sub-test 'intelligence' quotient scale, in the same procedure that Wechsler used for translating the full scale weighted score into a full scale intelligence quotient. It was intended in that procedure that the mean standard scores for each of the age groups be given an IQ value of 100, the zero point set at 10 P. E. below the mean, and IQs 90, and 110, be set at one P. E. below and one P. E. above the mean respectively. These values are considered measures of individual variability, which refers to the differences among individuals in a single trait, (*Table 1, col. 4*).

III. To establish the extent of trait variability, which refers to the variability from trait to trait within the individual the following procedure is suggested attendant upon the findings of Hull [2], and Pieron [3], and evaluated by Anastasi [1]. They respectively estimate trait variability to be 80% and 75% as great as individual variability. The individual standard scores, which relate to the individual weighted scores, (*Table 1, col. 3, and 1 respectively*) and which establish the individual sub-test 'intelligence' quotients, (*col. 4*) are now recalculated so that each standard deviation used in the denominators will be 77.5% (an average) of the original subtest age group standard deviations. This now permits the measure of deviation (a ratio of sub-test intelligence quotient divided by the average of the ten sub-test intelligence quotients, which is 100 for the standardization population) to be evaluated statistically. Since the

average of the ten sub-tests intelligence quotients is 100 for the standardization population, the values for these ratios may be read off from (*Table 1, col. 4*) to evaluate the specific patient's scores against the norms. Thus by finding ratio value from (*col. 4*) and by interpreting the extent of deviation by making probability statements of a strictly statistical nature from the accompanying standard scores (*col. 5*) reflective of trait variability, reliable deviation ratios may be evaluated.

Table 1. SHOWING THE METHOD OF ESTABLISHING NORMS FOR SCATTER ANALYSIS USING AS AN EXAMPLE THE OBJECT ASSEMBLY SUB-TEST, 20-24 YEAR OLD GROUP (MEAN = 10.4; S.D. = 2.93).

1	2	3	4	5
wgt. sc.	X-M	z sc. a	sub-test I.Q.*	(z sc., $\frac{\text{X-M}}{2.27}$)
1	-9.40	-3.21	52	
2	-8.40	-2.87	57	
3	-7.40	-2.52	63	
4	-6.40	-2.18	68	
5	-5.40	-1.84	73	-2.38
6	-4.40	-1.50	78	-1.94
7	-3.40	-1.16	83	-1.50
8	-2.40	- .82	88	-1.06
9	-1.40	- .48	93	
10	- .40	- .14	98	
11	+ .60	+ .20	103	
12	+1.60	+ .55	108	
13	+2.60	+ .89	113	+1.15
14	+3.60	+1.23	118	+1.59
15	+4.60	+1.57	123	+2.03
16	+5.60	+1.91	128	+2.47
17	+6.60	+2.25	133	

*I.Q. $= \dfrac{6.745 + \text{z sc. a}}{6.745}$ $\qquad$ $2.27 = 77.5\%$ of 2.93

The use of ratios would seem to be justified because the zero, though not absolute, will be stable enough so that error resulting from its use will not be greater than that inherent in the test, and will not influence the results noticeably.

SUMMARY AND CONCLUSIONS

A method is outlined to achieve comparability of scores on the Wechsler sub-tests, and to evaluate their dispersion. The advantages of this method are: (a) its preciseness, (b) the advantages of a ratio scale, for the different ranges of age and IQ, and (c) use as a training method. The writer believes that its optimal value will come from its use in conjunction with studies of intra-test variability. The deviation norms have been calculated for the Wechsler Scale for Children, in addition to those for Form I for Adolescants and Adults, and are available on request from the author.

REFERENCES

1. ANASTASI, A. and FOLEY, J. *Differential Psychology.* (Rev. Ed.), New York: Macmillan, 1949.
2. HULL, C. L. Variability in the amount of different traits possessed by the individual. *J. Educ. Psychol.*, 1939, *23*, 436-439.
3. PIERON, H. L'heterogenite normale des aptitudes. *Annee Psychol.*, 1940-41, *41-42*, 1-13.
4. WECHSLER, D. *Measurement of Adult Intelligence.* Baltimore: Williams and Wilkens, 1944.

EDITORIAL OPINION

THE INFLUENCE OF YALE UNIVERSITY ON AMERICAN PSYCHOLOGY

During the last five years, your editor has received more correspondence from readers on the topic of helping students to evaluate graduate training programs than on any other subject. A large per cent of this correspondence has reflected dissatisfaction on the part of students and their advisors with the shortcomings of various graduate training programs, and also with the difficulty of obtaining reliable information concerning what any one department has to offer. In pondering over the question concerning how to convey information to students concerning the characteristics of an ideal training program, it occurred to us that it might be valuable to cite the example set by a graduate department which has been outstandingly successful in all the services rendered to students. In our opinion, the influence of the Yale department on American psychology has been greater than that of any other university during the period from 1925 to 1950. It was during this period that Yale turned out the large number of leaders who have gained prominence in American psychology; the AMERICAN PSYCHOLOGICAL ASSOCIATION has had for its presidents in 1948, 1949 and 1951 three recent graduates of the Yale department. In searching for the reasons for this solid achievement, one looks first to the quality and quantity of the teaching staff which included at one time ARNOLD GESELL, CLARK L. HULL, MARK A. MAY, WALTER R. and CATHERINE C. MILES, and last but not least, ROBERT M. YERKES. These people could work together as evidenced by the cooperative interprise of the *Institute of Human Relations*. To a large extent, the compartmentalization of men with different interests which exists at other universities was broken down. Psychology at Yale was able to hold its own and obtain its just due in competition with other departments in the University. Sufficient funds were secured to support extensive research programs and to provide assistance for all deserving students. Best of all was the morale of the students who showed an "esprit de corps" which has lasted beyond student days. One had the feeling that here was a place where important things were being done. The atmosphere was stimulating, friendly and conducive to academic accomplishment. Out of this came Hull's learning theory, Gesell's child development studies, and Yerke's comparative researches all of which rate as among the major contributions of the period.

Much of the credit for producing such an ideal situation must go to the senior staff members who established the pattern. One of their outstanding characteristics was the personal interest shown in individual students. These men were interested not only in their teaching obligations but also in promoting the advancement of their students both academically and professionally. We recall with pleasure the seminars held by Dr. Hull in which scholarly interest was combined with healthy conviviality, and which led for a time to exchange seminars and baseball games with the Columbia department. Dr. Gesell was a leader in foreseeing that clinical material must be brought into the university in large enough quantities to satisfy research and clinical needs. To do this, the University had to establish clinics within its own control from which the well-known child development studies came. Dr. Yerkes is remembered by the force which his own personality had on students, and for the efforts made on their behalf both while they were students and later in advancing their professional careers. Dr. Yerkes was consistently positive in his liking for people, and in the things which he often gratuitously did for them. It seems to us that the Yale department has set a pattern which others might emulate. Here was a program which was genuinely *well rounded* not only in its academic aspect but also in relation to the personal services offered to students. At a time when many departments are understaffed, undersubsidized, compartmentalized, divided within them-

selves, commercially minded, narrow in theoretical orientation, unbalanced in representation of various "schools", and poor in morale of staff and students, it behooves the promising young graduate student to make thorough investigations of all curricular and extracurricular aspects of graduate training programs. There are too many instances of promising students becoming frustrated and developing unhealthy personality reactions because they do not get what they want. The APA has a Committee on Training in Clinical Psychology whose findings should properly belong to the members of the association. Let us hope that individual findings will be made available to those who are qualified to evaluate them.

F. C. T.

VALUE JUDGMENTS IN CLINICAL PSYCHOLOGY

The professional operations carried on by clinical psychologists do not have much social significance until they lead to value judgments upon which actions and decisions can be based. As would be expected, individual practices vary along a range from the most nondirective methods at one extreme to the most directive techniques at the other. Just at present, many clinical psychologists make every effort not to use their own value judgments to influence clients. In support of their viewpoint, they utilized all sorts of evidence ranging from exhortation and appeals to democratic principles to citations of research reports. Without entering into controversy concerning the relative merits of directive or nondirective methods, it seems important to comment upon the professional responsibilities of scientists in relation to value judgments. First, it is obvious that clinical psychologists should use value judgments wherever there is valid indication. Second, these value judgments should be advanced tentatively and nondogmatically, granting as much responsibility as possible to the client for their acceptance or rejection. Third, the clinical psychologist should not assume professional infallibility but should anticipate and compensate for inevitable errors in his judgment. Fourth, the clinical psychologist need not develop anxiety over the possibility of errors of ommission or commission. Legal medicine has long since established the principle that an individual practitioner can only be held responsible for operating according to the standard (average) qualifications existing at *time and place*. The personal equation is recognized as operating in all human activities. To err in judgment is human, and all that can be expected is for you to do the best you can. Clinical psychologists should not be timid in advancing value judgments particularly when supported by a consensus of scientific opinion. It may be just as much of an error not to make value judgments when indicated, as to make invalid judgments unsupported by scientific evidence or common sense.

F. C. T.

CALL FOR PAPERS

As soon as can be arranged, this journal plans to begin publication of a series of symposia known as CLINICS consisting of groups of papers describing standard practices and case handling methods currently in use in various installations about the country. These clinics will be patterned after such medical publications as *Surgical Clinics of North America* in which groups from various cities report their experiences using new methods. In the same manner as successive medical symposia might compare the results of using x ray and surgery in the treatment of tumors, so in the field of clinical psychology it would be appropriate to compare experiences with the results of various methods of counseling and psychotherapy with different clinical groups. Your editor hereby issues an invitation for interested groups to submit proposals for the publication of symposia relating to important problems in any field of clinical psychology. Priority will be assigned to projects which present new developments in clinical practice or which involve more extensive research data

than has heretofore been available. Although it will be expected that all contributions shall be condensed as much as possible, a limited amount of space will be made available for verbatim transcriptions of interviews particularly when unusual methods or results are presented. We are not particularly interested in publishing articles which merely report further experiences with methods which are already well-known and which therefore do not make any original contribution. It is hoped that we will be able to publish at least one or two symposia in this field every year.

Research suggestions. It seems important to construct a battery of tests in the field of biometrics which could be used to estimate the rate of aging in the different tissues of the organism as a clue to life expectancy. Among the recognized signs of aging are graying or baldness of the hair, skin changes including cellular pigmentation and fatty infiltration, decreased tissue elasticity, joint disorders, and various neurological signs including slower reaction time and the development of tremors. Unfortunately there are no objective indices for many of these signs. While thinking about these problems, we have made the casual observation that tissue elasticity may be evaluated by a simple test using the skin on the back of the hand. If the skin is grasped firmly and elevated from ¼ to ½ inch, great differences will be observed in the speed with which it snaps back into its normal flat position. It should be simple to develop a device which would record the speed and shape of the curve of elasticity. Assuming that the aging of the tissues of the organism in general will also be reflected in aging of the central nervous system, it should be possible to correlate the appearance of physical signs of senility with evidences of declining mental function. The syndrome of *premature senility* is occasionally seen as early as the second, third or fourth decades of life, while in more fortunate persons there may be few signs of senility into very old age. The rate of tissue aging, and the factors which control it, deserve more intensive study and consideration in research planning.

F. C. T.

LETTER TO THE EDITOR

To the Editor:

With respect to the Monograph Supplement CURRENT MENTAL HYGIENE PRACTICE published in the January 1951 issue of the Journal, it should be noted that Harvey J. Kramer and Roger A. Bricker contributed valuable assistance in the collection of the data on which it was based. Due to an unfortunate oversight, their names were omitted and their efforts did not receive proper acknowledgment.

Yours truly,

ROBERT TYSON, *Hunter College*

February 27, 1951.

BOOK REVIEWS

ROGERS, CARL R. *Client-centered Therapy.* Boston: Houghton-Mifflin, 1951, pp. 560. $4.00.

With this book Dr. Rogers traces the evolution of nondirectivism during the past ten years from the status of being primarily a method of counseling and psychotherapy to that of becoming a basic philosophy for human relations. It was during this period that systematic foundations for nondirectivism were synthesized depending heavily on Self-psychology and phenomenonological psychology. This book expounds a thoroughgoing philosophy for all types of clinical practice and human relations, and any critical evaluation must not overlook the fact that much of the material is *philosophical* rather than scientific in its origin. For this reviewer, both of Dr. Rogers' books have been extraordinarily difficult to review objectively because they present a confusing amalgamation of clinical intuitions and research findings upon which is superimposed a theoretical and interpretive system of very questionable validity. If the content of *Counseling and Psychotherapy* and *Client-centered Behavior* be analysed statement by statement, it will be discovered that their scientific value is constantly besmirched by a large number of *non sequiturs* reflecting Dr. Rogers' own personal and theoretical biases for and against different methods. On a grammatical level, one discovers a constant use of superlatives when referring to the values of his own methods and the use of depreciatory inflections when referring to other methods. When trying to read these works objectively, one has to constantly guard against a subtle salesmanship which is quite convincing. Dr. Rogers is an enthusiast with more than a little inclination to evangelize. To us, this proselyting attitude considerably diminishes the value of any scientific presentation because of the need for constantly guarding against propagandizing.

On more serious levels, it appears that much of the philosophical foundation constructed in support of nondirectivism is obsolete and has already been largely abandoned elsewhere. In his reliance upon "Growth principles" and the hypothesis that "that the individual has a sufficient capacity to deal constructively with all those aspects of his life which can potentially come into conscious awareness", (p. 24) Dr. Rogers appears to adopt a purposive, teleological vitalism in which "Good" is always expected to triumph over "Evil" in a somewhat Pollyannaish manner. While *normal* people do usually possess sufficient "growth" resources to solve their problems with a minimum of interference, it may be argued that such people *are not really sick* in the sense that by its very nature psychological illness is characterized by a relative inability to solve one's own problems caused by a breakdown of resources. Let us speak no longer of vague "growth resources". If such resources exist, they can be identified objectively, manipulated therapeutically, and strengthened using all available methods eclectically. Nondirectivism has no corner on this market.

But the main difficulty comes with Dr. Rogers' attempt to use the psychology of the Self and phenomenological principles as panaceas for all personality problems. It cannot be denied that phenomenological principles provide the most valid explanations for *some* categories of personality problems, and here there can be no quarrel with their use. There is no question that the use of nondirective methods can be extended into many previously untouched areas. But they are definitely not the answer in all areas of human relations. The world has been treated to a horrible example of the failure of nondirectivism to curtail such pathological forces as Hitlerism. Rigid attitudinal and ideational systems such as are found in paranoid states and other attitudinal disorders are more apt to be *strengthened* rather than ameliorated by nondirective handling. One of Freud's most important contributions was the insight that pathological attitudinal systems tend to defend themselves through

the mechanisms of resistance and repression. We must conclude that in their present form the philosophical foundations of nondirectivism are grossly inadequate for the tasks which Dr. Rogers expects them to fulfill.

On the positive side, this book makes a great contribution in presenting refinements of nondirective methods and in extending the indications for their use. In so far as the argument of this book stands upon objective research data, it very materially advances our understanding of the nature and applications of nondirective methods.

PLANT, JAMES S. *The Envelope.* New York: Commonwealth Fund, 1950, pp. 299. $3.00.

Rare is the book which presents a new and distinctive organization of theory and facts which goes beyond the mere compilation of the ideas of others. The late Dr. Plant was respected and loved as one of the pioneers of the child guidance movement in America. Although not as well known for literary contributions as others in the field, he was recognized as one of the most original thinkers in the field. Serving for twenty four years before his death in 1947 as Director of the Essex County (N. J.) Juvenile Clinic, Dr. Plant had a rich experience involving more than ten thousand problem children. On the basis of this wide experience, twenty-one areas of childhood "problems" are intensively discussed. In every area, Dr. Plant has been able to add something beyond the ideas of those who passed before him. This is a genuinely eclectic book, one in which every type of psychological influence is evaluated and given proper weight. It is required reading for all interested in the problems of children.

HELLERSBERG, ELIZABETH F. *The Individual's Relation to Reality in Our Culture.* American Lecture Series. Springfield, Ill.: Thomas, 1950, pp. 128. $3.25.

The Horn-Hellersberg test is a drawing completion test which has been adapted from Horn's *Art Aptitude Inventory* for the purpose of objectively evaluating a person's relation to reality (the psychiatric concept of orientation). This monograph outlines the rationale of the test, describes its administration, and cites illustrative cases with samples of typical results. The author claims to identify several signs which may be taken as indicators of a breakdown of contact and adaptibility to reality. These signs probably have suggestive significance but there is no objective data in this monograph in support of the author's intuitions.

YACORZYNSKI, G. K. *Medical Psychology.* New York: Ronald, 1951, pp. 535. $6.00.

In our opinion, this is the first textbook of medical psychology which is solidly based on the theory and data of modern experimental psychology. Its basic orientation is psychobiological, depending heavily on behaviorism and psychophysiology and in which field it is completely adequate. In the areas of anthropology, social psychology and group dynamics, the discussions are less adequate involving only 38 pages out of 500 in the text proper. This heavy overweighting in the somatic tradition is probably not a desirable characteristic in 1951 since it means that medical students will be exposed to a viewpoint which, although needed, is not currently being given so much emphasis in the field of psychology proper.

WOLFF, WERNER. *The Threshold of the Abnormal.* New York: Hermitage, 1950, pp. 473. $6.50.

This is a rather elementary textbook of abnormal psychology with added chapters presenting a brief survey of methods of diagnosis and psychotherapy. The presentation is on a college level and gives a well rounded but superficial orientation to all factors contributing to abnormal behavior. In some respects this is the most modern text in abnormal psychology that we have seen.

WEPMAN, JOSEPH M. *Recovery from Aphasia.* New York: Ronald Press, 1951, pp. 276. $4.50.

The author is Clinical Instructor in Otolaryngology (Speech Pathology) and Lecturer in Psychology at the University of Chicago. The book is intended as a textbook discussing the nature, diagnosis and treatment of the aphasias. Because of the author's dual background in psychology and medicine, the presentation is well-rounded and authoritative.

ROBINSON, FRANCIS P. *Principles and Procedures in Student Counseling.* New York: Harper, 1950, pp. 321. $3.00.

Dr. Robinson is Professor of Psychology at Ohio State University. His theoretical orientation may be regarded as being experimental-scientific and eclectic in its approach. Without actively sponsoring any particular technique, methods are presented with supporting research data relating to their indications and contra-indications. The level of discussions is rather elementary, being directed toward personnel workers in all fields. The treatment makes up in solidness and conservatism what it lacks in originality and imaginativeness.

LORAND, SANDOR (Ed.) *The Yearbook of Psychoanalysis. Vol. VI.* New York: International Universities Press, 1951, pp. 307. $7.50.

The sixth edition of the Yearbook continues the previous pattern of presenting a collection of psychoanalytic papers which have been judged as being of special significance.

LORAND, SANDOR. *Clinical Studies in Psychoanalysis.* New York: International Universities Press, 1950, pp. 272. $4.00.

This book consists of a series of essays and clinical notes on various aspects of psychoanalysis.

BIRREN, FABER. *Color Psychology and Color Therapy.* New York: McGraw-Hill, 1950, pp. 284. $4.50

This is a popular presentation of some of the highlights of the history and present status of the study of the influence of color on human life. In style it is a well-written and interesting compilation of scientific and empiric observations. Its greatest value will be for the layman who wishes to know something about color psychology.

FISCHER, SIEGRIED. *Principles of General Psychopathology.* New York: Philosophical Library, 1950, pp. 327. $4.75.

The author was formerly Professor of Psychiatry and Neurology at the University of Breslau. The book consists of a series of factual essays describing the principal phenomena of abnormal psychology largely from an outdated organic viewpoint.

ISLAND PRESS ANNOUNCES PUBLICATION OF A NEW AND ENLARGED EDITION OF

SELF CONSISTENCY: A Theory of Personality

by Prescott Lecky

Edited and interpreted with Addition of New Materials

by Frederick C. Thorne, M.D., Ph.D.

This new edition contains more than a hundred pages of previously unpublished papers and research data documenting Mr. Lecky's theory of integration as the primary personality goal. The abridged first edition contained the basic outline of self-consistency theory but omitted the supporting research evidence which is here published for the first time. Several new interpretive chapters written by Dr. Thorne are intended to extend the practical implications of the theory.

"A highly stimulating volume for the student of personality."—*Peabody Journal of Education.*

"Rich in both critical ideas and creative concepts."—*Psychological Bulletin.*

PRICE: $3.50 **Order from**

ISLAND PRESS

470 West 24th Street New York 11, N. Y.

PRINCIPLES OF PERSONALITY COUNSELING

by FREDERICK C. THORNE, M.D.

University of Vermont

This new textbook presents a new and eclectic system of Personality Counseling, including detailed discussions and case presentations of all known methods.

PRICE: $5.00. ORDER NOW

THE MONOGRAPH SERIES

No. 1.	Billingslea: Bender Gestalt Test	$ .90
No. 2.	Carter & Bowles: Qualitative Aspects of Psychological Examining	$1.25
No. 3.	Training in Clinical Psychology	$1.50
No. 4.	Critique of Nondirective Therapy	$1.25
No. 5.	Buck: The H-T-P Manual	$2.00
No. 6.	Karpman: Objective Psychotherapy	$2.00
No. 7.	Statistics for the Clinician	$2.00
No. 8.	Tyson: Current Mental Hygiene Practice	$2.00

ORDER TODAY

Journal of Clinical Psychology

BRANDON, VERMONT

CLINICAL STAFF

MEDICAL STAFF OF PENNSYLVANIA

Leslie R. Angus, M.D.
Robert Devereux, M.D.
Ruth E. Duffy, M.D.
Herbert H. Herskovitz, M.D.
Ruth Stephenson, M.D.

PSYCHOLOGICAL STAFF OF PENNSYLVANIA

Edgar A. Doll, Ph.D.
Director of Research
Michael B. Dunn, A.M.
Robert G. Ferguson, A.M.
Edward L. French, Ph.D.
Mary J. Pawling, A.M.

MEDICAL STAFF OF CALIFORNIA

Charles M. Campbell, Jr., M.D.
Director of Clinical Services
Richard H. Lambert, M.D.
Consulting Psychiatrist
Ivan A. McGuire, M.D.
Consulting Psychiatrist
David L. Reeves, M.D.
Consulting Neurologist

Thomas W. Jefferson, Ph.D.
Clinical Psychologist

MATURITY
through Education with Therapy

Never before has the need been so great for maturity in our adult population . . . for clear, objective thinking and a realistic facing-up to problems---both personal and social. The emotionally disturbed child stands a poor chance of becoming a mature and valuable citizen unless he has specialized help . . . understanding guidance and encouragement in solving the problems commensurate to his age.

When, in your practice, you feel that a school-age patient needs specialized education and guidance, we invite you to let us evaluate the potential outcome. When the intelligence is normal, but emotional disturbances block the ability to learn, our experienced staff will carefully study the details and offer a considered report.

Please address your inquiries to
John M. Barclay, Registrar

DEVEREUX FOUNDATION
CALIFORNIA ..... PENNSYLVANIA

Devereux Schools

HELENA T. DEVEREUX, Director
SANTA BARBARA, CALIFORNIA :-: DEVON, PENNSYLVANIA